

VOLUME 24    NUMBER 3    SEPTEMBER 1961

# SOCIOMETRY

A JOURNAL OF  
RESEARCH IN SOCIAL PSYCHOLOGY

## *Articles by*

BORGATTA

LUCKY

WEST

MESSICK

NETTLER

HAMBLIN, MILLER and WIGGINS

SCHUESSLER

Published by

THE AMERICAN SOCIOLOGICAL ASSOCIATION

# SOCIOMETRY

## EDITOR

JOHN A. CLAUSEN

## EDITOR-ELECT

RALPH H. TURNER

## ASSOCIATE EDITORS

EDGAR F. BORGATTA  
ORVILLE G. BRIM, JR.  
LAUNOR F. CARTER  
JAMES S. COLEMAN  
ERVING GOFFMAN

SIDNEY SEIGEL  
SHIRLEY A. STAR  
WILLIAM A. SCOTT  
GUY E. SWANSON  
ROBERT F. WINCH

## EDITORIAL CONSULTANTS

WENDELL BELL	THEODORE M. MILLS
LEONARD S. COTTRELL, JR.	FREDERICK MOSTELLER
FRED E. FIEDLER	ROBERT RADLOW
ALLEN R. HOLMBERG	THEODORE R. SARBIN
ALEX INKELES	EARL S. SCHAEFER
HERBERT C. KELMAN	JOHN THIBAUT
MELVIN L. KOHN	EUGENE A. WEINSTEIN
WILLIAM W. LAMBERT	MARION RADKE YARROW

## EXECUTIVE OFFICER

ROBERT BIERSTEDT

## ADMINISTRATIVE OFFICER

JANICE H. HOPPER

*Published four times a year*

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL ASSOCIATION

at The Boyd Printing Co., Inc., 49 Sheridan Ave., Albany 10, N. Y.

*Sociometry was founded in 1937 by J. L. Moreno*

Beginning October 1, 1961, all new manuscripts should be addressed to the Editor-elect, Dr. Ralph H. Turner, Sociometry, Department of Anthropology and Sociology, University of California, Los Angeles 24, California. Business communications should be addressed to Janice H. Hopper, Administrative Officer, American Sociological Association, New York University, Washington Square, New York 3, New York.

Copyright © 1961 by THE AMERICAN SOCIOLOGICAL ASSOCIATION

Entered at the Post Office at Albany, N. Y., as second-class matter.

# SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 24 • NUMBER 3 • SEPTEMBER 1961

## CONTENTS

Editor's Note	217
Role-Playing Specification, Personality and Performance <i>Edward F. Borgatta</i>	218
Perceptual Congruence of Self and Family Concepts as Related to Marital Interaction..... <i>Eleanore Braun Luckey</i>	234
Class Origin of Scientists..... <i>S. Stewart West</i>	251
The Perceived Structure of Political Relationships... <i>Samuel Messick</i>	270
Good Men, Bad Men, and the Perception of Reality.. <i>Gwynn Nettler</i>	279
Group Morale and Competence of the Leader <i>Robert L. Hamblin, Keith Miller and James A. Wiggins</i>	295
A Note on Statistical Significance of Scalogram.... <i>Karl F. Schuessler</i>	312

## SOCIOMETRY

### EDITORIAL POLICY

*Sociometry* is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup, and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

---

Annual subscription rate \$9.00; special rate to members of THE AMERICAN SOCIOLOGICAL ASSOCIATION \$4.50. Single issues \$2.50.

*Specifications for Manuscript Preparation for Sociometry* may be obtained from the Editorial Office.

*Back Issues* dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of *change of subscriber's address* should be sent to THE ASSOCIATION four weeks in advance, indicating the old as well as the new address.



## Editor's Note

The editors of *Sociometry* are repeatedly called upon to assess not only the significance and scientific adequacy of manuscripts but also their appropriateness for the journal. Using even a very broad definition of social psychology, a number of the manuscripts received do not come primarily under this heading. Some are concerned with aspects of social organization not highly relevant to the psychological realm; others report psychological research better qualified by adjectives other than *social*; still others are methodological analyses relevant to a wide range of behavioral science studies.

Not surprisingly, the editors vary in their views of "relevance to social psychology" just as they vary in their assessments of adequacy and significance. On the whole, we are committed to flexibility in representing the field as it is accessible to us, whether in respect of subject matter, of methodology or of article length. For this reason, no "departments" or "sections" within the journal have seemed necessary or desirable. Yet frequently our editors are disposed to suggest that a given paper might be appropriate as a research note, but not as a full article. Or they comment that the ideas or data presented in a particular manuscript are of sufficient interest to social psychologists to merit publication in *Sociometry* in brief form if not in an extended version.

One such manuscript appears in the present issue—that by Karl Schuessler. While the paper is concerned primarily with a statistical issue, it seemed to our referees germane to the interests of many readers of *Sociometry*. At the same time, we felt that a quite brief presentation was called for, and we are delighted that Dr. Schuessler was willing substantially to abridge his original manuscript.

We should like to encourage the submission of other brief papers. Our field needs replications of many studies which generalize beyond immediate research contexts. Such replications often do not require elaborations of the theoretical rationale or the basic methodology; they therefore lend themselves to concise presentation. Again, brief reports of fragmentary data, when these have clear significance for the field, or brief methodological notes would be welcomed.

We do not know what the effect of such a call for brief manuscripts will be, but we shall try to provide an outlet for good "research notes." To avoid any misunderstanding, this does not constitute a change in *Sociometry's* policy of considering quite lengthy manuscripts if these carry their weight in contribution to knowledge.

## Role-Playing Specification, Personality, and Performance<sup>1</sup>

EDGAR F. BORGATTA, *University of Wisconsin*

The research reported in this paper involves the following: Persons of known personality characteristics are asked to play assigned roles that involve characteristics designed to be (a) similar and (b) different from their own. The report presents data on regularities associated with differences of personality characteristics, with differences of assigned role specifications, and also with the *congruency* between the previously measured personality characteristics and the assigned role specifications. Emphasis is placed on the specific personality measures used in the study. The effects that are observed are described through interaction process scores and through assessments, made immediately after each role playing situation, on how well each person played and how much each person enjoyed the assigned role.

### SELECTION OF PERSONALITY CHARACTERISTICS

Previous analysis indicates that much of the variance of scores based on peer judgments can be accounted for by a few concepts, particularly assertiveness, sociability, intelligence, and emotionality (1, 3, 4, 5). Further, a broad range of peer judgments has been studied in a variety of contexts, and the structure of such judgments has been replicated (1). For these reasons, it appeared desirable to use as personality characteristics in this study those derived from peer judgments and defined consistently in the replicated analysis of the structure of the judgments. The central concepts for analysis as personality characteristics in this study were two composite scores for the subjects derived from peer rankings after a five-man discussion session lasting 80 minutes and covering four topics for which group decisions were requested. Subjects were students at New York University, generally in the sophomore year at initial contact. The use of four independent topics permitted exposure of the individual in different contexts and in randomly distributed alliances, and thus, presumably, more broadly based peer judgments. Each composite peer score was the sum of peer rankings on four trait or behavior characteristic descriptions: *Assertiveness* was based on (a) Is very active, (b) Takes initiative, (c) Makes many suggestions, and (d) Shows assertiveness. *Emo-*

<sup>1</sup> This study stems from the research program, "The Variables and Conditions of Small Group Interaction," at New York University. This research was supported in whole or in part by the U. S. Air Force under contract AF 49 (638)—195, monitored by the Air Force Office of Scientific Research of the Air Research and Development Command (AFOSR TN 60-682).

*tionality* was based on (a) Is very tense, (b) Gets upset easily, (c) Tends to be very nervous, and (d) Emotionality. In the analyses of the peer judgments from which these data are drawn, the items included in the composite scores tend to be relatively pure in the two factors involved. The composite scores of Assertiveness and Emotionality were relatively independent: e.g., in an analysis of 155 male subjects from which the current samples were drawn, the correlation was .10. Data are also presented for analyses based on parallel attitudinal scales.

#### RESEARCH DESIGN

On the basis of the prior Assertiveness score, the subjects were ordered in three units of approximately equal size. The principle of ordering was not revealed to the observers on the project until after the final stages of data collection for the entire project. The subjects, of course, had no knowledge of this ordering.

Approximately one year after the data for the composite scores were available, a series of three-man groups was constituted for role-playing sessions. With minor exception, each group was composed of the combination of High, Middle and Low (H, M, and L) on prior Assertiveness score, and each group was composed so that members had no prior experimental contact in discussion groups. For this stage of analysis 29 groups or 87 subjects were recruited.

Each role-playing session involved three scenes. Each scene involved three persons with assigned roles to portray, and the directions were to carry on the role-playing for 20 minutes. The assigned roles were identified as Person 1, Person 2, and Person 3 in the directions; subjects who were H, M, and L were rotated through these identifications systematically by selection prior to the sessions. The order of the three scenes, however, was held constant.

Each scene involved status position identifications plus descriptions of characteristics considered to be consistent with the position. For example:

Person 1 is a boy, 12 years old, who tends to be somewhat withdrawn. He is not particularly active, doesn't talk very much, and doesn't seem to mix well with others his age.

Person 2 is the boy's mother, who is emotional, generally tense, and gets upset easily.

Person 3 is the boy's father, who is an active and directive person, almost authoritarian in his manner.

The scene is in the living room of the family's home. *You* are that family. You are discussing the boy's problems and the future.

In each of the scenes the persons described were designed to require (a) Assertive, (b) Submissive, and (c) Emotional performance. The role-playing specifications (assigned roles) were arranged so that each subject would

perform under each type of requirement as he proceeded through the three scenes. Subjects appeared to respond to the instructions for role-playing without difficulty. It had been anticipated from prior experience that subjects would have no particular difficulties in taking the roles of the opposite sex, and this appeared to be the case here. Some indication of the ease with which the role-playing was undertaken may be seen in the high level of self-ratings on enjoyment discussed later in the paper.

The role-playing sessions were observed using a revision of Bales's scoring categories (3), here called Interaction Process Scores (IPS's). Table 2 presents a brief identification of the categories. In addition, after each role-playing scene, the participants completed a single-sheet rating form dealing primarily with how well each person performed and how much each person enjoyed playing the assigned role. Two observers also rated the role-players after each scene on the same questions.

When this stage of the data collection was completed, the identical design was repeated, using three additional but comparable scenes to provide a replication. The groups were composed on the same basis, again under the requirement of no prior experimental contact among the three-man group members. In this case, 24 groups or 72 of the 87 subjects constituted the sample size.

#### ANALYSIS OF THE DATA

Some of the findings from this study may be seen directly in Table 1. Series A, Series B, and Series A and B are the study, the replication, and the combined data.<sup>2</sup>

Note the following: Persons selected as H, M, and L in Assertiveness one year previously have on the average Total IPS's that correspond to this order not only generally, but also for each role specification, and this is true both in Series A and B. In other words, independent of the assigned role, Total IPS is a function of Assertiveness score.

That the assigned roles have differential meaning for the subjects is obvious in the consistent ordering of the roles by Total IPS as Assertive, Emotional, and Submissive.

That differential response is possible for subjects in these assigned roles is obvious in the substantially higher Total IPS's of the Low Assertive persons in the Assertive and Emotional assigned roles than of the High Assertive persons in the Submissive roles.

Note one additional incidental replicated finding in the comparison of Series A and B. The Total IPS's for the All column of Series B are higher,

<sup>2</sup> Data for Series A and B are separately reported only for Table 1 to illustrate the parallel results.

TABLE 1

*Total Interaction Process Score by Prior Peer Assertiveness Score and Role Specification \**

	Assertiveness Score	Role Specification			
		Assertive	Submissive	Emotional	All
Series A	H	216.3	114.4	190.0	173.6
	M	208.8	96.3	172.2	159.1
	L	168.1	91.4	140.6	133.4
Series B	H	284.8	106.9	246.3	212.7
	M	255.0	91.0	223.1	189.7
	L	235.7	84.7	180.4	166.9
Series A+B	H	247.3	110.9	215.5	190.6
	M	229.7	93.9	195.3	173.0
	L	198.7	88.4	158.6	148.6

\* Series A, 29 groups or 87 subjects; Series B, 24 groups or 72 subjects. Participation of a subject in each group was with persons with whom he never participated previously. High, Medium and Low Assertiveness scores were based on peer rankings in five-person discussion groups one year before role-playing. High, Medium and Low scores were defined by approximately equal thirds. The total interaction process scores are for 20 minutes in discussion (mean number of responses).

indicating a facilitation effect with practice. The increase arises only in the Assertive and Emotional assigned roles, and if anything, there is indicated a decreased rate for the Submissive assigned role. This might appear inconsistent, but this is what should be expected by facilitation when the instructions are for Submissive performance.

Aside from the variation noticeable on the Total IPS's, when Total IPS is controlled by converting the profile of separate IPS's into percentage rates, a number of qualitative differences are found. The largest differences are associated with the assigned roles (Table 2). For example, for the Series A and B in the Submissive assigned role (all subjects) only about 12 per cent of the total responses are IPS 17, *Shows antagonism*, in contrast to 28 per cent for the Assertive assigned role and 33 per cent for the Emotional assigned role. About 5 per cent of the Submissive assigned role responses are IPS 18, *Ego defensive*, compared to 3 per cent for the Emotional and 2 per cent for the Assertive assigned roles. About 17 per cent of the Submissive assigned role responses are IPS 15, *Shows inadequacy*, compared to 10 per cent for the Emotional and 4 per cent for the Assertive assigned roles. The profile for the Emotional assigned role is particularly characterized by high percentages of response in IPS 14, *Disagrees*, and IPS 17, *Shows antagonism*, while that of the Assertive assigned role has a high percentage of responses in IPS 06, *Procedural suggestion*, in addition to differences already noted.





Qualitative differences also occurred according to prior Assertiveness score. For example, the Low Assertiveness persons showed a greater proportion of agreement, both active and passive (IPS 04 and 05), asked for fewer opinions (IPS 13), showed less antagonism (IPS 17), but showed relatively more disagreement (IPS 14).

In Table 3, data are presented on the ratings of how well each person played and how much he enjoyed playing the assigned role. For the self-ratings, there is only a slight tendency for H, M, and L persons on prior Assertiveness scores to be ordered on either criterion. The self-ratings appear to be at a uniformly high level compared to peer and observer ratings, the only appreciable deviation being that H, M, and L persons all tended to rate enjoyment of the Submissive assigned role low.

TABLE 3

*Ratings on Role-Playing Ability and Enjoyment by Prior Peer Assertiveness Score and Role Specification (Series A + B) \**

HOW WELL DO YOU FEEL EACH PERSON PLAYED THE PARTICULAR ROLE THAT WAS ASSIGNED?

	Assertiveness Score	Role Specification			
		Assertive	Submissive	Emotional	All
Observer rating	H	1.94	1.79	1.81	1.85
	M	2.25	1.60	1.83	1.89
	L	.64	1.51	.89	1.01
Peer rating	H	1.94	2.23	2.04	2.07
	M	2.53	1.30	1.81	1.88
	L	1.36	1.40	1.21	1.32
Self-rating	H	2.87	2.91	2.89	2.87
	M	2.74	2.72	2.87	2.77
	L	2.57	2.68	2.64	2.63

HOW WELL DO YOU FEEL EACH PERSON ENJOYED PLAYING THE ROLE ASSIGNED?

Observer rating	H	2.17	1.04	2.41	1.87
	M	2.36	.81	2.17	1.78
	L	1.43	.87	1.58	1.30
Peer rating	H	2.06	1.11	1.75	1.64
	M	2.26	.77	1.87	1.64
	L	1.55	.87	1.34	1.25
Self-rating	H	2.81	2.34	2.91	2.69
	M	2.72	2.24	2.70	2.55
	L	2.79	2.34	2.55	2.56

\* Questions were answered after each 20-minute role-playing session. Range of ratings was from 0-4.



Large differences are seen for H, M, and L persons, however, in the peer and observer ratings. While the ratings appear ordered on H, M, and L, the striking difference is in low ratings for Low Assertiveness persons. Low Assertiveness persons, however, get *relatively* high ratings on role-playing ability in the Submissive assigned role. This corresponds to the expectation that congruence of personality and assigned role should produce superior performance.

Still, in spite of the fact that Low Assertiveness persons are seen to play the Submissive assigned role best of the three by both peers and observers, these raters view this as the least enjoyable assigned role for the Low Assertiveness persons.

While the ratings by peers and observers are both high for role-playing ability and enjoyment for the High Assertiveness persons, these are not uniformly the highest ratings. A notion of congruency appears reasonable here, but obviously something else is operating. The Medium Assertiveness persons get the highest ratings in the Assertive assigned role, which suggests that they take that assigned role well and comfortably. If they take the Assertive assigned role well, it may be they take others well also. For this to be the case with the Submissive assigned role, they should simulate Low Assertiveness persons in that role. In terms of Total IPS (Table 2), they appear to do so. In other words, a possibility arises that a Medium rating on Assertiveness signifies persons who are as Assertive as the assigned role or moment requires, rather than usually, rarely, or half assertive. Such a person should enter these roles with relative ease—the versatile person. That this is the fact is at least partially supported in the qualitative profiles of IPS's (Table 2), where it is seen that Medium persons have the lowest tension rates generally (IPS 15). A question arises as to why the role-playing ability of the Medium persons is not higher for the peer and observer ratings of the Submissive assigned role. Here, speculatively, the interpretation of the rating task comes under scrutiny. The behavior of a Submissive person may contradict notions of competence in interpersonal relations, and thus behavior that simulates submission effectively may be negatively rated on more general grounds than role-playing ability. Effective simulation is indicated here, since the pattern of ratings for the Medium persons on the Submissive assigned role are similar to those for the Low persons. The interpretation suggested is that the Medium persons are being rated as though they were Low persons.

Interpretation of ratings in the Emotional assigned role for H, M, and L persons is not as direct, and only a general suggestion is made here. Assertiveness and Emotionality have a slight positive correlation as indicated previously. The ordering of ratings appears consistent with this for High and Low

persons, with the Medium persons closer to the High than the Low. Since the notion of submissiveness implies withdrawal or inability to be assertive, and neither of these appear consistent with displaying emotionality, the ordering appears appropriate on this *ad hoc* basis. On the speculative side for the Medium persons, the ability to take the Emotional assigned role is as high as for the High persons, consistent with a notion of versatility. If meaning is attributed to the minor differences in higher ratings for the High persons, elaboration of the notion of versatility as associated with competence may be pursued. Competence is generally associated with rationality, and Emotionality is composed primarily of "irrational" aspects of behavior. Presumably, the versatile person would be less comfortable in the irrational role.

#### *Prior Peer Emotionality Scores*

On the basis of prior peer Emotionality scores, the subjects were placed into three approximately equal groups. For the Series A and B data presented, the number of observations were 54 High, 57 Medium, and 48 Low Emotionality persons. In terms of combinations occurring in the groups, the distribution was within expectations of random assignment, which means that the lowest frequency of expected participation for any person was with two persons of his own class, i.e., combinations of HHH, MMM, or LLL. For all practical purposes, the distribution permits a concept of viewing H, M, and L persons on prior peer Emotionality scores as participating with equivalent panels of co-participants, under control of the distribution of the (relatively independent) peer Assertiveness scores. The implications of observation are somewhat different here than with the peer Assertiveness scores, where essentially each person was observed in his relative modal position on Assertiveness according to prior measures.

Contrasting the profiles of the High and Low in Table 4, High persons on prior peer Emotionality scores show more *antagonism* (IPS 17), more *tension* (IPS 16), and less *disagreement* (IPS 14). The differences in interaction rates are relatively small, with the High persons having the highest rate. Low persons appear to be characterized by higher proportions of activity in categories associated with task attention (IPS 06 and IPS 12).

In Table 5, data are presented on the ratings of observers, peers, and self. There is no apparent tendency for the self-ratings to be ordered according to prior Emotionality scores for the role-playing ability question, while for the role-playing enjoyment question there is a somewhat smaller amount of enjoyment indicated for Low Emotionality persons. Most of this small differential is contributed by the lower self-rating in the Emotional assigned role, a difference which is compatible with a notion of assigned role congruency.

TABLE 4  
*IPS Percentage Profiles by Prior Peer Emotionality Score by Role Specification (Series A + B)*

Prior Emotionality Score Role Specifications	High				Medium				Low			
	A	S	E	Total	A	S	E	Total	A	S	E	Total
Number of Cases	54	54	54	162	57	57	57	171	48	48	48	144
IPS 1 Social acknow.	.1	.1	.1	.1	.1	0.0	.1	.1	.1	.1	0.0	.1
IPS 2 Shows solidarity	1.3	1.2	.9	1.1	1.3	1.1	.8	1.1	1.3	1.7	1.3	1.4
IPS 3 Laughs	.2	.6	.3	.3	.7	.8	.7	.8	.2	.8	.2	.3
IPS 4 Acknowledges	.6	.9	.4	.6	.7	.9	.7	.7	.9	.5	.5	.7
IPS 5 Shows agreement	.6	1.7	.6	.8	.5	1.2	.9	.8	.6	1.9	.8	.9
IPS 6 Procedural sugg.	6.1	3.9	3.6	4.7	6.0	3.1	3.2	4.4	8.1	4.9	3.6	5.9
IPS 7 Suggests solution	2.5	.5	1.2	1.6	2.8	.8	.8	1.7	2.3	.8	1.2	1.6
IPS 8 Gives evaluation	36.1	37.0	25.7	32.3	34.4	34.4	27.6	31.9	36.5	31.4	27.1	32.2
IPS 9 Self-analysis	0.0	.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.2	0.0	0.0
IPS 10 Redirected agg.	.1	.2	.3	.2	.5	.5	.3	.4	.1	.6	.4	.3
IPS 11 Gives information	1.6	3.4	1.1	1.8	1.2	2.1	1.1	1.4	1.4	1.8	.9	1.3
IPS 12 Draws attention	5.3	7.8	4.2	5.4	4.4	8.0	5.2	5.4	6.9	6.2	5.3	6.2
IPS 13 Asks evaluation	1.2	.3	.7	.8	1.2	.4	.5	.8	1.2	.3	.7	.9
IPS 14 Disagrees	8.1	10.1	12.0	10.0	10.4	11.8	12.9	11.6	10.9	11.0	14.6	12.3
IPS 15 Shows inadequacy	4.6	17.8	9.6	9.1	3.5	16.3	10.9	8.6	3.4	17.1	8.6	7.9
IPS 16 Tension increase	.1	.3	.2	.2	.3	.3	.2	.2	.1	.3	.1	.2
IPS 17 Shows antagonism	29.4	9.6	36.0	28.1	29.5	12.2	30.7	26.8	23.4	14.8	31.3	24.5
IPS 18 Ego defensiveness	2.3	4.6	3.1	3.0	2.3	6.0	3.4	3.4	2.4	5.5	3.4	3.3
Total	100.2	100.1	100.0	100.1	99.8	99.9	100.0	100.1	99.8	99.9	100.0	100.0
Mean number responses for 20 min. session	232.4	107.0	211.2	183.8	215.8	88.8	177.4	160.7	228.5	97.3	180.8	168.9

TABLE 5

*Ratings on Role-Playing Ability and Enjoyment by Prior Peer Emotionality Score and Role Specification (Series A + B) \**

HOW WELL DO YOU FEEL EACH PERSON PLAYED THE PARTICULAR ROLE THAT WAS ASSIGNED?

	Emotionality Score	Role Specification			
		Assertive	Submissive	Emotional	All
Observer rating	H	1.98	1.57	2.07	1.88
	M	1.28	1.58	1.07	1.31
	L	1.58	1.78	1.40	1.58
Peer rating	H	2.18	1.83	2.09	2.04
	M	1.70	1.46	1.42	1.53
	L	1.96	1.65	1.54	1.72
Self-rating	H	2.76	2.78	2.81	2.78
	M	2.68	2.74	2.77	2.73
	L	2.69	2.79	2.81	2.76
HOW WELL DO YOU FEEL EACH PERSON ENJOYED PLAYING THE ROLE ASSIGNED?					
Observer rating	H	2.26	1.20	2.65	2.04
	M	1.93	.61	1.68	1.41
	L	1.75	.92	1.83	1.50
Peer rating	H	2.26	1.37	2.05	1.90
	M	1.86	.68	1.60	1.38
	L	1.73	.69	1.27	1.23
Self-rating	H	2.87	2.26	2.85	2.64
	M	2.81	2.40	2.81	2.67
	L	2.63	2.25	2.46	2.44

\* Questions were answered after each 20-minute role-playing session. Range of ratings was from 0-4.

Both in terms of observer and peer ratings, High persons on prior peer Emotionality scores are rated highest on role-playing ability (All column), and the differential would be even more striking if it were not for the ratings in the Submissive assigned role. Here the differences are less regular and smaller. It should be noted that the Low persons are not lowest in the ratings; rather, the Medium persons are lowest. The high ratings of High persons by observers and peers are most emphasized in the Emotional assigned role, and this is as should be according to a notion of assigned role congruency. Speculatively, what is to account for the apparent high ratings of the Low persons relative to the Medium persons? As such, Emotionality is not a positively valued characteristic in society. Being low on Emotionality ratings might result either from being withdrawn, or being controlled, phlegmatic,

or rational. The latter qualities are more favorably seen in society, and it may be that performance by such persons may lead to general, favorable ratings in this case, rather than to ratings as directed on the criterion of how well the roles were played. On the other hand, it may be that Low persons on Emotionality are better able to portray characteristics of others. The latter suggestion, however, raises other questions. For example, why should the High persons be rated higher on the Assertiveness assigned role? It is not *explanation* to say that the ratings are a reflection of a general ability. Speculatively, in the case of the high ratings in the Assertiveness assigned role for High persons on prior peer Emotionality scores, a possible explanation might arise in the kinds of behavior that lead to high Emotionality scores that may also relate to favorably rated role-playing ability (such as personal involvement), or in the noticed small dependency between Assertiveness and Emotionality scores.

From the point of view of enjoyment ratings by peers and observers, High persons on prior peer Emotionality score clearly are seen to enjoy the role-playing more than Medium or Low persons in the three roles. The differences in ratings for the Medium and Low persons are not clearly interpretable. In the Assertive assigned role, Low persons are rated to have the lowest enjoyment, but in the Emotional assigned role observers rate the Low persons higher than the Medium persons, while the peers do the opposite. What is of interest in these data is that the peers rate the enjoyment of the Low persons low while not correspondingly rating the role-playing ability low. Here it should be recalled that Low persons tend to rate themselves relatively low on enjoyment of role-playing, particularly in the Emotional assigned role.

#### RE-ANALYSIS UTILIZING ATTITUDINAL MEASURES

From prior testing (15-18 months earlier), scores were available on personality inventory scales that are conceptually related to peer Assertiveness and Emotionality scores. Those selected for examination in back analyses (similar to that of Emotionality above) were from the Guilford-Zimmerman Temperament Survey, GZ 3 Ascendancy and GZ 5 Emotional Stability. The correlation of the peer Assertiveness score with GZ 3 Ascendancy in the panel of subjects from which the current samples were drawn was .33, and for peer Emotionality and GZ 5 Emotional Stability the correlation was -.16.

When persons were grouped as H, M, and L according to the GZ 3 Ascendancy score, no substantial difference in mean interaction rate was found, and the differences in profiles appeared to be smaller and less consistent than for peer Assertiveness. More *antagonistic behavior* (IPS 17) was noticed for High persons on GZ 3 Ascendancy, and slight differentials of *less tension* (IPS 15) and *disagreement* (IPS 14) also occurred.

In Table 6, the rating data are presented for the GZ 3 Ascendance H, M, and L persons. Self-ratings appear to have the same pattern as for peer Assertiveness. The overall differences between High and Low persons appear to be slightly smaller in the case of GZ 3 Ascendance, but clearly the trend is for High persons to be rated high on role-playing ability, as was the case with peer Assertiveness. In pattern, enjoyment ratings appear to correspond to the ratings of role-playing ability. In comparison to performance in the Assertive and Emotional assigned roles, observer ratings are highest on role-playing ability for Low persons in the Submissive assigned role. The data are not as clear cut for the peer ratings, however. Of some interest is the relatively low rating on role-playing ability for High persons made by the observers,

TABLE 6

*Ratings on Role-Playing Ability and Enjoyment by Prior G-Z 3 Ascendance Score and Role Specification (Series A + B) \**

HOW WELL DO YOU FEEL EACH PERSON PLAYED THE PARTICULAR ROLE THAT WAS ASSIGNED?

	G-Z 3 Ascendance Score	Role Specification			
		Assertive	Submissive	Emotional	All
Observer rating	H	2.07	1.41	1.91	1.80
	M	1.48	1.90	1.50	1.67
	L	1.25	1.62	1.09	1.32
Peer rating	H	2.34	1.79	1.93	2.02
	M	1.70	1.76	1.76	1.74
	L	1.75	1.38	1.36	1.50
Self-rating	H	2.86	2.88	2.95	2.89
	M	2.84	2.86	2.96	2.89
	L	2.43	2.57	2.50	2.50

HOW WELL DO YOU FEEL EACH PERSON ENJOYED PLAYING THE ROLE ASSIGNED?

Observer rating	H	2.33	.75	2.48	1.82
	M	1.90	1.24	1.96	1.71
	L	1.81	.74	1.70	1.42
Peer rating	H	2.16	1.02	1.89	1.69
	M	2.12	1.04	1.64	1.60
	L	1.68	.70	1.42	1.23
Self-rating	H	2.88	2.32	2.91	2.70
	M	2.76	2.30	2.72	2.59
	L	2.68	2.30	2.51	2.50

\* Questions were answered after each 20-minute role-playing session. Range of ratings was from 0-4.

with a corresponding low enjoyment rating. Speculatively, the assigned role congruency notion is suggested in reverse (i.e., persons out of roles consistent with their personality perform poorly).

In the analysis of H, M, and L persons on the GZ 5 Emotional Stability scale, no substantial differences in interaction rate occur. Some interesting qualitative differences do occur, however. Consistent with the peer Emotionality analysis, GZ 5 Emotional Stability is associated with less tense behavior (IPS 15) and more disagreement (IPS 14), and somewhat more behavior associated with attention to task (IPS 06 and IPS 12). Contrary to the prior findings, however, High persons on the GZ 5 Emotional Stability scale are high on antagonistic behavior (IPS 17).

TABLE 7

*Ratings on Role-Playing Ability and Enjoyment by Prior G-Z 5 Emotionality Score and Role Specification (Series A + B) \**

HOW WELL DO YOU FEEL EACH PERSON PLAYED THE PARTICULAR ROLE THAT WAS ASSIGNED?

	G-Z 5 Emotionality Score	Role Specification			
		Assertive	Submissive	Emotional	All
Observer rating	H	1.92	1.55	1.47	1.64
	M	1.46	1.76	1.51	1.57
	L	1.69	1.46	1.54	1.76
Peer rating	H	2.42	1.86	1.78	2.02
	M	1.89	1.57	1.81	1.73
	L	1.79	1.59	1.33	1.57
Self-rating	H	2.72	2.75	2.72	2.73
	M	2.69	2.82	2.89	2.80
	L	2.74	2.67	2.67	2.69

HOW WELL DO YOU FEEL EACH PERSON ENJOYED PLAYING THE ROLE ASSIGNED?

Observer rating	H	2.19	.86	2.05	1.77
	M	1.87	.81	1.96	1.55
	L	2.05	1.15	2.08	1.76
Peer rating	H	2.39	.94	1.92	1.75
	M	1.87	.76	1.70	1.44
	L	1.74	1.23	1.31	1.43
Self-rating	H	2.83	2.58	2.72	2.71
	M	2.61	2.24	2.79	2.58
	L	2.85	2.21	2.56	2.54

\* Questions were answered after each 20-minute role-playing session. Range of ratings was from 0-4.



In Table 7, the rating data are presented for the H, M, and L persons on the GZ 5 Emotional Stability scale. In the self-ratings, only a slight association of High persons with high rating is indicated. The observer ratings do not present an interpretable panel of findings, but the H, M, and L are ordered for the peer ratings both on the Assertive and Emotional assigned roles for role-playing ability and enjoyment. Interpretation at this point becomes difficult, since GZ 3 Ascendancy and GZ 5 Emotional Stability are correlated at .36 among the subjects from which these samples are drawn, and peer Assertiveness is correlated to GZ 5 Emotional Stability at .18.

#### SUMMARY AND DISCUSSION

In this analysis attention has focused on the concomitants of characteristics of individuals in their rated ability, enjoyment, and performance in specified role-playing situations. Two role descriptions were designed to correspond to the assertiveness vs. non-assertiveness (submissiveness) dimension, and a third to emotionality. Persons were ordered on the basis of prior peer Assertiveness scores, and groups with the combination H, M, and L were systematically placed in each assigned role position of three fixed role-playing situations. After each of the three sessions, rating forms were completed by the participants and the observers, providing observer, peer, and self-ratings on role-playing ability and enjoyment. The interaction was scored on a revised set of observation categories (IPS's). Data were collected for two series involving the same subjects in non-overlapping groups. Analysis of the results appears to confirm a notion of facilitation in the situation of personality and assigned role congruency, but there is reason to suggest a need for attention to other factors. For example, favorable ratings may be related to societal approval of characteristics rather than to the experimental instructions of rating how well the assigned role-playing was done. When one plays an assigned role, he may be rated favorably because he is mistaken for someone with the favorable characteristics rather than because he has done a good job of role-playing.

Through the availability of other prior scores, the analysis was extended to examine findings for a classification of persons on a criterion of peer Emotionality scores. With this relatively independent analysis, additional support for facilitation of personality and assigned role congruency was found.

Additional re-analyses on the basis of two attitudinal scale scores (GZ 3 Ascendancy and GZ 5 Emotional Stability) were presented as parallels to the analyses based on prior peer scores. Some power to reproduce the findings of peer Assertiveness is demonstrated for GZ 3 Ascendancy. Differences that occur are not directly interpretable, and inevitably would appear to lead the investigator to a re-analysis of the development of the attitudinal scales.

These types of difficulties were those anticipated and proposed to emphasize the appropriateness of utilizing peer (and observer) assessments as an important direct criterion basis for the classification of persons.

The common view of the relationship between personality and social structure includes two important propositions. (a) Personality types tend to gravitate to positions with consistent normative demands. (b) Occupancy of positions with given normative demands tends to alter the personality towards consistency with the normative demands. Stated in mundane language, an individual tends to do things consistent with his personality, and what he does tends to affect his personality. Concomitant with these propositions is a notion that when persons of given characteristics are placed in situations that make demands of them consistent with their characteristics, they will be able to satisfy the demands well and will do so with relatively little discomfort or with relatively more enjoyment. Implicitly, this notion would underlie the above two propositions. In this research we have not tested the notion of congruency of personality characteristics with role demands directly, but we have at least moved into the vicinity of such an exploration. It must be remarked that the concept of role is itself one that must be viewed as primitive and more subject to intuitive verbal gymnastics than to use in the study of personality and social structure in a formal sense (2). Of two obvious ways in which "role demands" could be manipulated, (a) by creating particular structures or (b) by instructing persons to assume them, we have chosen the latter. Within the scope of the types of positions identified (father, mother, son, policeman), the former is not subject to easy manipulation. To the extent that behavioral characteristics described are consistent with the position mentioned, the assigned roles should be interpretable, and the ability to assume them is of central concern to this research. While there is no doubt a level of detachment involved in role-playing (It's not for real!), there is also a level of ambiguity about the reality of structures created synthetically in the laboratory. These are limitations on the research in this area, but they are not total disqualifiers. Additional approaches should be taken in this study of personality and social structure, however, including longitudinal regression analyses of selective processes—particularly with emphasis on the comparison of persons of given characteristics who do and do not follow modal patterns in selection of roles.

*Manuscript received: June 15, 1960*

*Revised Manuscript received: March 21, 1961*

Edgar F. Borgatta  
Department of Sociology  
University of Wisconsin  
Madison, Wisconsin

## BIBLIOGRAPHY

1. Borgatta, E. F., "Rankings and Self-Assessments: Some Behavioral Characteristics Replication Studies," *Journal of Social Psychology*, 1960, 52, 279-307.
2. Borgatta, E. F., "Role and Reference Group Theory," in L. S. Kogan (Ed.), *Social Science Theory and Social Work Research*, New York: National Association of Social Workers, 1960.
3. Borgatta, E. F., "A Systematic Study of Interaction Process Scores, Peer and Self Assessments, Personality and Other Variables," *Genetic Psychology Monographs* (in press).
4. Borgatta, E. F., L. S. Cottrell, Jr., and J. H. Mann, "The Spectrum of Individual Interaction Characteristics: An Inter-Dimensional Analysis," *Psychological Reports, Monograph Supplement 4*, 1958, 4, 279-319.
5. Carter, L. F., "Recording and Evaluating the Performance of Individuals as Members of Small Groups," *Personnel Psychology*, 1954, 7, 477-484.

## Perceptual Congruence of Self and Family Concepts as Related to Marital Interaction<sup>1</sup>

ELEANORE BRAUN LUCKEY, *University of Connecticut*

A recent series of studies investigated the relation of marital satisfaction to congruent perception of certain self and other concepts (9, 10, 11). Congruent perception was defined in terms of a score indicating the discrepancy between two concepts as described by the subject's marking of items on an adjective check list. Five pairs of concepts were investigated: (a) *self* and *self as marked by spouse*; (b) *self* and *ideal self*; (c) *self* and *parent of the same sex*; (d) *spouse* and *parent of the opposite sex*; and (e) *spouse* and *ideal self*. It was anticipated that in each instance the group of subjects scoring high in marital satisfaction would also be those holding reliably more congruent perceptions. The findings generally supported this positive relationship, but also permitted the formulation and answering of other challenging questions: When perceptual differences exist between the maritally satisfied and unsatisfied, what are the dimensions of these differences? Can these be described in terms of personality traits? Are these differences related to the sex of the perceiver? And what, if any, meaning might these findings contribute to a theory of interpersonal relations as they exist in the most intimate and intense of all human relationships—marriage? These are the questions the present study proposes to investigate.

It is important to note at the outset that this study is concerned with the degree of agreement of certain perceived personality concepts held by two groups of married subjects who differ reliably in their degree of marital satisfaction. It does not deal with the trait differences between satisfied and less satisfied subjects; it does *not* attempt for example, to answer such questions as: Does the wife who is satisfied with her marriage see her spouse as more loving than the less satisfied wife sees hers? Each concept is always considered in relation to a second concept: self in relation to the self as perceived by the spouse; self in relation to one's ideal self, etc.

### THE SAMPLE AND INSTRUMENT

The two samples investigated were designated as "satisfactorily" and "less satisfactorily" married on the basis of the highest and lowest couple scores on

<sup>1</sup> This article combines papers presented at the Groves Conference on Family and Marriage, The State University of Ohio, Columbus, 1960, and the Midwest Sociological Society Annual Conference, St. Louis, Missouri, 1960. The analysis of the data was made while the author was on the faculty of the University of Iowa and was supported by the Iowa Child Welfare Research Station. Special appreciation is expressed to A. R. Mangus for his valuable critical evaluation of the study.

23 items of the Locke (7, 8) and Terman (15) marital scales. Forty-one "satisfactorily married" (S) couples and 40 "less satisfactorily married" (LS) couples were chosen by several phases of follow-up from an original group of 594 former students of the University of Minnesota, 454 of whom responded to the initial follow-up to assess marital satisfaction. All those finally classified in the two extreme groups—well satisfied and dissatisfied—participated in the further stages of the research.<sup>2</sup> The two samples were found to be homogeneous in regard to 26 items of descriptive personal information such as age, number of years married, number of children, income, social class, residence, education, vocation, religion, and a variety of parental background factors.

The Interpersonal Check List (5, 6) was used to describe concepts of self, spouse, mother, father, and ideal self; it was chosen primarily on the basis of its construction and the selection and arrangement of items. It is important to note that the ICL is not used in this study as a diagnostic instrument or for personality classification. These are the uses for which it was designed, and the scoring methods which Leary has developed are in keeping with this diagnostic purpose. In this study, however, the instrument was used as an organized and uniform list of descriptive statements about self and others, and an appropriate scoring method was devised.

The level of measurement employed in this study has been designated by Leary as "Level II, Conscious Description." There is only one criterion for determining Level II data: conscious verbal report by the subject. Directions to the subject instruct him to mark each item he considers to be generally descriptive of himself at the present time, and to proceed in the same way to describe the other persons indicated. Scores thus obtained are not to be confused with what are known as "assumed similarity scores." AS scores are secured by directing the subject to mark the items for himself and *as he thought the other would mark them for himself*, in other words to mark them as if he were in the other's shoes (1, 3).

The check list provides an organized way in which to deal with what a subject *says* about himself and others. When used as a classification of direct, surface statements that a subject makes about himself and others, the check list may be considered a valid expression of the way he chooses to present himself and his view of the world. On this basis the ICL can be said to have content validity.

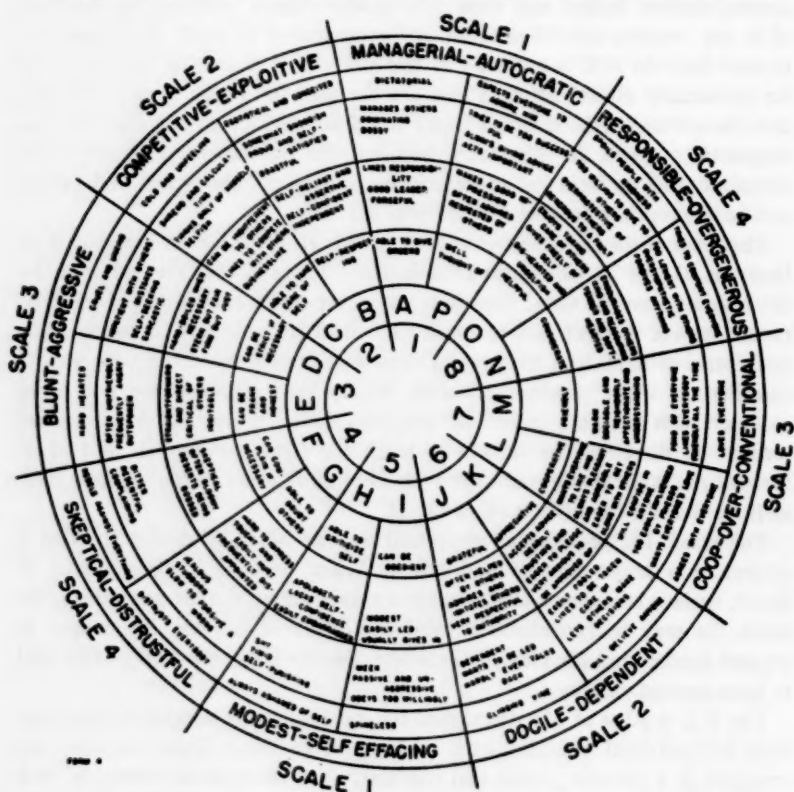
The ICL is a list of 128 descriptive self-referent items selected to represent eight interpersonal variables, each made up of 16 items. These variables are arranged in a circular profile and combined into descriptive octants in such a way that the opposite octant represents an opposing variable; i.e., each

<sup>2</sup> For more complete description of the sample and criterion instruments, see reference 9.

intersecting line on the graph may be thought of as a continuum with the subject's score being located according to the number of items checked within that specific category. The average test-retest reliability of the ICL has been reported by Leary (6) to range from  $+.75$  to  $+.83$  (average  $+.78$ ) when correlations were done on each octant. As important as reliability measures are the intervariable correlations. Adjacent variables on the circular continuum are more closely related than non-adjacent, and the relationship between two variables is a monotonic decreasing function of their separation (5). The averages of the correlations for variables one step apart decrease as more

FIGURE 1

*Interpersonal Check List Illustrating the Classification of Interpersonal Behaviors*



\* Adapted from Figure 1, *Multilevel Measurement of Interpersonal Behavior*, p. 2 (6).  
Permission for reproduction granted by Timothy Leary.



distant variables are correlated, thus opposite quadrants may be considered opposite ends of the same continuum.

For the purpose of this study the two octants making up the one continuum have been designated as a scale, of which there are four for each concept measured. Scale 1 is made up of variables termed Managerial-Autocratic, Octant AP, at one end of the continuum, and Modest-Self-effacing, Octant HI, at the other. Scale 2, Competitive-Exploitive, BC, vs. Docile-Dependent, JK; Scale 3, Blunt-Aggressive, DE, vs. Cooperative-Over-conventional, LM; Scale 4, Skeptical-Distrustful, FG, vs. Responsible-Overgenerous, NO (see Figure 1). Some of the items are: Octant AP, dictatorial, forceful, acts important, able to give orders; Octant HI, timid, apologetic, able to criticize self; BC, cold and unfeeling, businesslike, self-respecting; JK, clinging vine, easily fooled, eager to please, grateful.

#### SCORING

Octant scores were determined by the number of items checked in each specific octant. Scale scores were then obtained by subtracting the smaller octant score from the larger; so that, if a subject marked eight items in Octant AP and four items in Octant HI, his scale score would be 4 in the AP octant. This score would indicate that he saw the personality that he was describing as being "managerial and autocratic," as contrasted with the opposing octant "modest and self effacing."

When the degree of congruency between two concepts was tested for significant differences between the S and LS groups, a discrepancy score was determined by taking the absolute difference of the scale scores on the two concepts being compared. These differences were then tested for statistical significance by means of the *t* test. These data are presented in Tables 1 and 2.

Because the ICL is conveniently organized in terms of eight descriptive and related personality variables, it is possible to define discrepancies not only by their magnitude but in terms of whether one concept has more or less of the personality qualities described by Scales 1, 2, 3, and 4.

In order to observe these qualitative aspects of the personality concepts that were being compared, the mean scale scores for the S and the LS groups were plotted on the usual ICL circular profile (Figures 2 and 3). Shaded portions of the circle represent the items marked by the two groups when comparisons were made on the pairs of concepts: (Circle I) self and self by spouse; (II) self and ideal self; (III) self and parent of the same sex; (IV) spouse and parent of the opposite sex; (V) spouse and ideal. Because a scale score always falls in one or the other of the two octants that comprise the scale, shadings have been made only in the one octant in which the score lies.



TABLE 1

*Differences Between Less Satisfied and Satisfied Groups of Married Women When the Degree of Congruence of Concepts was Tested for Significance*

ICL Scale	Mean Discrepancy LS Group (N=40)	Mean Discrepancy S Group (N=41)	t
I. Self by wife: Wife by husband			
1	.568	.450	1.82*
2	.530	.489	.73
3	.539	.473	1.22
4	.584	.498	1.41
II. Self by wife: Ideal self by wife			
1	.638	.565	.97
2	.554	.493	.85
3	.565	.370	3.31**
4	.730	.508	4.04**
III. Self by wife: Mother by wife			
1	.735	.585	2.24*
2	.539	.484	.83
3	.545	.536	.15
4	.601	.599	.03
IV. Husband by wife: Father by wife			
1	.644	.515	1.77*
2	.599	.488	1.79*
3	.809	.577	3.41**
4	.709	.497	2.86**
V. Husband by wife: Ideal self by wife			
1	.602	.313	4.74**
2	.573	.392	2.97**
3	.724	.400	4.38**
4	.872	.439	6.37**

\* p .05.

\*\* p .01.

The shaded areas are proportioned so as to accurately represent quantitative scores obtained thus:

$$X \text{ Scale } 1 = \bar{x}_{AP} - \bar{x}_{HI}$$

$$X \text{ Scale } 2 = \bar{x}_{BC} - \bar{x}_{JK}$$

$$X \text{ Scale } 3 = \bar{x}_{DE} - \bar{x}_{LM}$$

$$X \text{ Scale } 4 = \bar{x}_{FG} - \bar{x}_{NO}$$

#### FINDINGS

Congruence of scores on the ICL is noted on the diagrams (Figures 2 and 3) by areas which are overlaid with two shadings. The distance between the outer edge of one shading and the outer edge of the other in the same octant

TABLE 2

*Differences Between Less Satisfied and Satisfied Groups of Married Men When the Degree of Congruence of Concepts was Tested for Significance*

ICL Scale	Mean Discrepancy LS Group (N=40)	Mean Discrepancy S Group (N=41)	t
I. Self by husband: Husband by wife			
1	.629	.427	3.61**
2	.566	.434	2.06*
3	.605	.441	2.73**
4	.557	.543	.22
II. Self by husband: Ideal self by husband			
1	.650	.379	3.99**
2	.497	.351	2.70**
3	.569	.451	1.64
4	.645	.474	2.44**
III. Self by husband: Father by husband			
1	.731	.550	2.55**
2	.558	.424	2.20*
3	.609	.550	.82
4	.670	.468	2.89**
IV. Wife by husband: Mother by husband			
1	.636	.521	1.46
2	.608	.545	.89
3	.665	.536	1.90*
4	.703	.534	2.56**
V. Wife by husband: Ideal self by husband			
1	.581	.469	1.72*
2	.572	.492	1.08
3	.564	.398	2.27*
4	.715	.437	4.15**

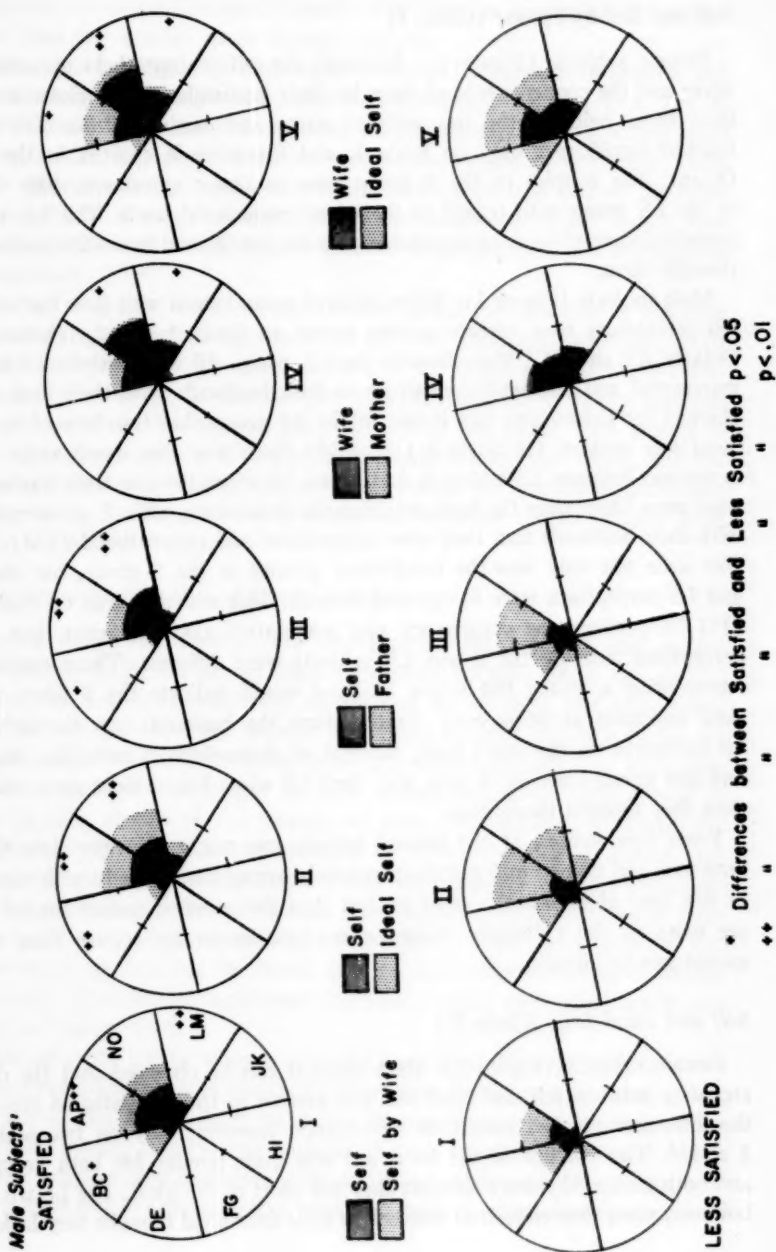
\* p .05.

\*\* p .01.

represents the difference in the mean scores between the two concepts under comparison. The plus marks (+) in the upper row of circles indicate the scales on which significant differences in the degree of congruence between the S and the LS groups were found in the previous studies. Those octants in which there are plus marks are those in which the satisfied subjects saw the concepts as reliably more congruent than the less satisfied subjects did; the S groups were found more congruent than the LS on every scale tested whether the difference reached statistical significance or not.



FIGURE 3. Interpersonal Check List Profiles indicating Mean Discrepancies between Satisfied and Less Satisfied Husbands when Concepts of Self, Spouse, Parents and Ideal Self are Compared



*Self and Self by Spouse (Circle I)*

Female subjects (Figure 2): Although the self-concepts held by satisfied wives and the concepts held of them by their husbands were in closer accord than those held by the less satisfied wives and husbands, the difference reached significance only on Scale 1, and the score is charted in the AP Octant. The couples in the S group were in closer agreement than those in the LS group with regard to the wives' managerial traits. The LS wives perceived themselves as being considerably less managerial than their husbands thought them.

Male subjects (Figure 3): Wives in the S group agreed with their husbands' self-perceptions to a reliably greater extent on Scales 1 and 2, (charted in Octants AP and BC) than those in the LS group. All wives attributed more managerial and competitive qualities to their husbands than their husbands claimed for themselves, but it was in the LS group that this lack of agreement was greater. On Scale 3 (DE-LM) there was also significantly less agreement between LS husbands and wives. LS wives thought their husbands were more blunt than the husbands thought themselves, while S wives agreed with their husbands that they were cooperative and conventional (LM). On this scale not only was the congruency greater in the S group, but the S and LS perceptions were in opposite octants. This was also true on Scale 4 (FG-NO) where the congruency was not statistically significant, but the perceptions held by the S and LS subjects were different. These qualities representing a strong but loving husband would indicate the S wives saw their husbands as being even "better" than the husbands saw themselves. LS husbands, on the other hand, thought of themselves as being less warm and less strong than the S men, and their LS wives found them even colder than they thought themselves.

From observations of the shaded octants one might also speculate that both men and women in S marriages not only agreed more closely with regard to the kind of person the other is, but that the satisfied spouse tended to see more of the culturally valued characteristics in his spouse than the spouse saw in himself.

*Self and Ideal Self (Circle II)*

Female subjects (Figure 2): Here again it can be observed that the discrepancy between self and ideal self was greater in the less satisfied group; the difference reached statistical significance, however, only on two scales, 3 and 4. The concept of the ideal self was quite similar for both groups, and both groups of women thought they fell short of the ideal. The LS wives, however, when they compared themselves with their ideal thought they lacked

generosity and cooperativeness (Octants NO, LM) to a significantly greater extent than the satisfied wives thought they did.

Male subjects (Figure 3): Although S men and LS men saw themselves and their ideals in the same relationship on all scales, on Scales 1, 2, and 4, S men came reliably closer to fulfilling their ideal. On Scale 3, where there was no statistically significant difference in the degree of congruity, there was a difference in the perceived traits. The LS men who desired the same warm qualities of cooperativeness (LM) and generosity (NO) that S men desired saw themselves with little or none of these qualities, and instead saw aggressiveness (DE).

We may conclude that for men in this sample, satisfaction in marriage was related to perceptual agreement of self and ideal self, and that, while satisfied men perceived the same qualities, albeit to a lesser extent, in themselves and their ideals, the less satisfied men saw the ideal as having qualities they did not have. For women the results were less clear; both groups of women perceived similar differences in themselves and their ideals, and the degree of congruency was significant in only two of the four scales. Both men and women in the LS groups saw themselves as falling shorter of their ideal selves than did the S subjects.

#### *Self and Parent of the Same Sex (Circle III)*

Female subjects: Although the discrepancies between self and mother were greater in the LS group, these differences between the S and the LS samples did not reach statistical significance on three of the four scales. The S women thought they differed from their mothers as much as the LS women thought they did except on Scale 1. S women thought they were more like their mothers with regard to their managerial traits; both groups of wives thought they were less managerial than their mothers. However, the direction of the difference was reversed on Scales 3 and 4. S women saw themselves as being more responsible (NO) and cooperative (LM) than their mothers, while the LS women thought their mothers exceeded them in these qualities. Thus if we assume that the mother is taken as a pattern, and somewhat as an ideal, for the female self, we can conclude that the S woman saw herself as an improvement over this ideal pattern while the LS wife saw herself as falling short.

Male subjects: Unlike women, men who were in satisfactory marriages saw themselves as reliably more like their fathers than the LS men on three scales: 1, 2, and 4. The men in the S group saw themselves as being generally much like their fathers; fathers were thought to be a bit more managerial (AP). The less satisfied men saw themselves as having similar qualities to those of their fathers but to a lesser extent except on Scale 4, where fathers were seen



as somewhat skeptical and distrustful (FG) but the subjects saw themselves as responsible and generous (NO).

Although marital satisfaction was not found to be associated with congruency of the female self-concept and concept of mother, it was associated positively with the male concepts. The direction of the differences between self and parent of the same sex indicates that LS women saw themselves as less autocratic and exploitive. It may be interpreted that LS subjects when comparing themselves with their same-sexed parents saw themselves as being less socially acceptable persons.

#### *Spouse and Parent of the Opposite Sex (Circle IV)*

Female subjects: Every one of the four scales indicated differences in congruence that were significant when women compared their husbands and fathers. Satisfied wives saw husbands and fathers as being alike; their husbands they thought even exceeded their fathers—who were obviously very good men, strong (Octants AP, BC) and warm (LM, NO). In contrast, LS wives thought their fathers were not very loving (LM, NO) and that their husbands were even less so. In addition, then, to there being significantly less similarity between LS husbands and fathers, there also appears to be considerable difference in the kinds of husbands and fathers L and LS women think they have.

Male subjects: When men compared their mothers and wives, the S men were found to see the two as alike to a statistically significant degree on Scales 3 and 4, but in a reverse relationship on Scales 1, 2, and 4. Men in S marriages saw their mothers as stronger (AP), more responsible (NO), and less docile (JK) than their wives, while men in LS marriages thought their mothers had less of the AP and NO qualities and more JK. Both S and LS men thought wives were more cooperative (LM) than mothers. Because the degree of congruency was significantly different on only two scales and at the .05 level, the relationship of congruent concepts of wife and mother to marital satisfaction is not clear. It appears, however, that male subjects in the S and LS groups perceived their womenfolk as being somewhat different kinds of women.

#### *Spouse and Ideal Self (Circle V)*

Female subjects: When the LS wives compared their self ideal with their perception of their husbands, four of the octants had no overlap whatsoever (LM, NO, DE, FG). On the other hand, the S women saw their husbands as having very much the same qualities as their ideal selves. The difference in congruence was significant at the .01 level on all four scales. The LS wives

who saw their ideal as being a predominantly responsible (NO) and generous (LM) person saw their spouses as having none of these qualities. On the contrary their spouses were thought to be distrustful (FG) and blunt (DE). Husbands of the LS women were obviously perceived as lacking in warm, loving qualities. The lack of agreement between the perceived husbands and ideal selves in the S group was primarily a matter of varying quantities within the same octants; i.e., husbands were seen as like the ideal, only less so.

Male subjects: Men in the S group perceived their ideals and their spouses as alike to a greater degree on Scales 1, 3, and 4 than LS men. On each of these scales as well as on Scale 2 the direction of differences is the same, with the ideal in each case having more of the quality than the spouse.

What may well be a sex difference is demonstrated on Scale 2, on which both S and LS men perceived their ideal as having a positive score in the BC octant; neither group of men saw their wives as positive scorers in this octant. The trait is one that is described as being "strong" or "dominant," and a perceptual difference of this kind would be in keeping with our cultural ideal which favors a greater degree of dominance in the male.

#### DISCUSSION

The implications of this study are numerous and far from crystal clear. On an over-all basis one is struck by the fact that the discrepancy scores of satisfied and less satisfied men differed more sharply than did those of women. Does this imply that perceptions held by the male are more closely related to marital satisfaction than those held by the female? Such an implication flies in the face of general "beliefs" about the marriage relationship. It has long been contended that marriage is chiefly the woman's interest and concern and is dominated by her attitudes and values.

Apparently the sharpest differences between S and LS subjects (both sexes) lay in their comparisons of the ideal self and spouse. The direction of these differences consistently indicated that LS subjects saw the ideal as greatly exceeding the spouse. S subjects, who thought the two were more similar, reflect the "idealization" of the spouse which has been found typical of the happy marriage.

Men who were satisfied with marriage came closer to fulfilling their ideal self-concept than less satisfied men, particularly on the dimensions of strength and warmth. Although both groups of men thought these were desirable qualities, the LS subjects saw themselves as lacking them. Whether this mirrors the husband's feelings of inadequacy within himself or is a response to his relationship with his wife would be mere speculation. In any case, these men apparently wished they were stronger and more affectionate than they believed themselves to be.

It was found that the relation of the perception of parents to marital satisfaction depended on the sex of the subject. Agreement of concepts of selves and mothers was not closely associated with marital satisfaction, but agreement of concepts of husbands and fathers was. Husbands were seen by S wives to exceed their fathers in desirable qualities. Men who were satisfied in marriage saw themselves as being like their fathers, with fathers being somewhat more managerial, but they did not see their wives and mothers as being any more alike than the LS men saw theirs. This seems to indicate that, for both men and women, it is important for the maritally satisfied to find husbands and fathers similar, but unimportant for both sexes to find wives and mothers alike. This suggests that the character of the male may be a more dominant factor in family relationships, both parental and spousal, than is the character of the female. It also would support the traditional view of the marriage relationship which holds that it is more fitting—hence more satisfying—for a wife to perceive her husband as a father and so relate to him as a subordinate, than it is for the husband to see and to relate to his wife as his mother. Beyond the sheer congruency of perception, it seems important to note that both male and female LS subjects seem to perceive their fathers as being cold, dominant personalities. Further studies need to be made in order to clarify these implications of the relation of parent concepts to marital adjustment.

Although it can in general be said that marital satisfaction was related positively to the congruence of certain self and other concepts, and that the quality of the perceptions was also related, no single meaningful generalization, which will apply to all these pairs of concepts and their association with marital interaction, can be inferred.

Actually, although all the discrepancy scores involve dyadic units (2), there are four specific sets of methodological issues involved in this group of five hypotheses, and each hypothesis has to be related to marital satisfaction in terms of its own individual content. In the first hypothesis there is only one object perceived, but there are two different perceivers, the spouses. In the second and third hypotheses the self which does the perceiving is at the same time one of the objects that is perceived. In the two remaining hypotheses, the subject perceives two objects outside himself. A fourth category of issues is involved in the second and fifth hypotheses which deal with the ideal self.

The positive relationship found among male subjects when their perceptions of themselves were compared with that which their wives held of them would tend to substantiate Mangus's (12, 13) theory that such congruent perceptions are associates of marital satisfaction. There is no ready explanation of why this should not be equally true for women subjects.

Whether the phenomenon of congruent perceptions held by spouses and marital satisfaction are causally related is a question to which no clue is given by this study. It may be theorized that the satisfaction of the marriage has given rise to the congruence, that wives who are happy with their marriages see husbands within the halo of their satisfaction and attribute to them the same—and more—good characteristics that the husbands claim for themselves. This and other studies (4) have indicated these factors are associated; what the precise nature of the association is, is yet to be discovered.

Although the congruency scores of this study are not the usual assumed similarity scores—and may or may not be the equivalent of such scores—they do share certain criticisms leveled at the AS and AD scores (2). These scores are all dyadic units in which one numerical score involves the combining of perceptions either *of* or *by* two or more persons (14, 2). Cronbach (2) has said such scores are “unparsimonious descriptions of events,” and holds that simpler units of measure may give less confusing results with less complicated statistical procedures. By analyzing scores of one concept, could we expect to discover relationships that are as meaningful as those that contain the two factors? Is it only the congruence of concepts which is meaningfully associated with marital satisfaction, or would a single measure of the trait qualities perceived *in* or *by* the subjects yield more easily interpretable links with satisfaction in marriage? Do we need only to be concerned with the expectation of the other as expressed through agreement of perception, or should we also be concerned with *what* the subjects see? For example, when self and self by spouse were compared, there were no statistically significant differences for either men or women on Scale 4 (charted in Octant NO). The individuals in both groups agreed with their spouses, yet what the S subjects saw in their spouses was something quite different from what the LS spouses saw in theirs, and what the S and LS subjects perceived in themselves was different. In like manner, congruency was not found related to marital satisfaction when the wife compared herself and her mother. It may be, however, that the perceptions that the S and LS women held of their mothers and/or themselves were reliably different.

Although congruency of the tested concepts has been found generally to be positively associated with satisfaction in marriage, the differences in the concepts themselves as they are perceived by S and LS people remain to be investigated.

Single scores, while sacrificing the aspects of interaction between two personalities suggested by a study in congruence, would permit a closer investigation of factors associated with marital satisfaction and each of the sexes. It is interesting to note that satisfied subjects of both sexes saw themselves

and both their parents as loving persons (Circles III and IV), but only the LS women saw themselves and mothers as loving; their fathers and spouses they thought were cold; the LS men saw their mothers and their wives as loving, but thought that they and their fathers were cold. This kind of perceptual difference between the two groups and between the two sexes leads to speculation as to whether this may express an attitudinal difference that is associated with marital satisfaction. The S subjects apparently saw loving qualities in both sexes; the LS saw women as warm but men as cold. Such findings may also indicate that the subjects who were able to make satisfactory marital adjustments were, indeed, those who were reared by two parents they saw as loving rather than by a warm mother married to a cold father.

This study also suggests that perhaps as important to building a theory of marital interaction as reducing dyadic scores to single scores is the necessity for investigating the relationship of more than two of the concepts simultaneously. For example, is there a different relationship between self, ideal, and parent of the same sex in the S and LS groups? A most challenging hypothesis is suggested by study of Circles I, II, III, and IV where self, spouse, and parents are compared. Do S and LS subjects perceive their parents as a spousal pair in a different relationship than they perceive themselves? The possibility of expanding the investigation to include other significant images should also be considered, for marital interaction probably involves a whole complex of personalities beyond the immediate ones of spouse, parents and ideal. However, at present there is no statistical method by which a complex of relationships of this kind can be satisfactorily analyzed, and if there were, the necessary interpretation of scores as attitudes associated with interpersonal behavior would be questionable.

This study has attempted to meet the criticism of the global score, usually employed as a measure of social perception (2), by observing the direction of congruency on four scales, each with a definable dimension. Although this technique undoubtedly makes possible a sharper definition of the dimension under observation, it necessitates a tremendous reliance on the Interpersonal Check List and its organization of personality characteristics. It is obvious to researchers in the area of perception that one of the first necessities for improved methods and techniques lies in the improvement of available instruments.

Undoubtedly the most meaningful contribution this study makes is in the multitude of perplexing problems it proposes and the hypotheses it suggests. These are problems and hypotheses which, when explored scientifically, will contribute to a developing theory of interpersonal relationships and, most specifically, of the spousal relationship.

## SUMMARY

Previous studies have indicated differences in the congruency of certain person perceptions as they relate to marital satisfaction. The present study has investigated these congruencies in terms of particular personality traits.

Couples were defined as satisfied (S) or less satisfied (LS) and were asked to mark their perceptions of (a) self, (b) spouse, (c) mother, (d) father, and (e) ideal on the Interpersonal Check List. This was then scored so as to yield four scale scores each describing a personality variable.

When the subject's self-concept was compared with the concept his spouse held of him, there were significant differences found between the S and LS male subjects. Self-concepts held by satisfied husbands agreed with their wives' perceptions of them more closely than perceptions of LS subjects; S wives saw not only the same traits in their husbands that their husbands saw in themselves but consistently saw their husbands as having more of each trait.

When the subject compared himself and his ideal, S men were found reliably more like their ideal selves on three of the four scales. In each case the ideal exceeded the self.

S men, when they compared themselves and their fathers, saw themselves as very like their fathers; LS men not only saw themselves as unlike their fathers but saw their fathers quite differently from the way S subjects saw theirs. LS men saw their fathers in relation to themselves as more dominant and more hostile.

When wives compared husbands and fathers, S women found that the two were significantly more similar to each other on all four scales and that husbands tended to exceed fathers on all traits. LS wives found husbands and fathers dissimilar, with husbands exceeding fathers on the dominant, hostile dimensions.

Both husbands and wives in the S groups found their spouses significantly more similar to their ideal selves than did the LS subjects. In each case where there were reliable differences, the S subjects saw their spouses as possessing the same traits as the ideal, but to a lesser extent. LS women, however, saw their ideals as warm and loving and their husbands as cold and skeptical. LS men saw wives as falling quite short of the ideal in the dimensions of autocracy and generosity.

These findings suggest that certain perceptual congruencies and dimensions on which the perceptions are congruent are related to the degree of satisfaction the subjects find in marriage. The importance of perceptual congruency and marital satisfaction seems related in some cases to the sex of the subject.

Possible interpretations and their contributions to a theory of marital inter-



action have been suggested, and important methodological issues have been raised. It has been made implicit that, before empirical research can hope to add much to the theory of marital interaction, improved instruments and techniques must be developed.

*Manuscript received July 18, 1960*

*Revised Manuscript received: March 22, 1961*

*Eleanore Braun Luckey*

*Department of Child Development and Family Relations*

*University of Connecticut*

*Storrs, Connecticut*

#### REFERENCES

1. Bronfenbrenner, U., "The Study of Identification Through Interpersonal Perception," in R. Tagiuri and L. Petrullo, *Person Perception and Interpersonal Behavior*, Stanford, California: Stanford University Press, 1958, 110-130.
2. Cronbach, L. J., "Proposals Leading to Analytic Treatment of Social Perception Scores," in R. Tagiuri and L. Petrullo, *Person Perception and Interpersonal Behavior*, Stanford, California: Stanford University Press, 1958, 353-379.
3. Fiedler, F. E., "The Psychological Distance Dimension in Interpersonal Relations," *Journal of Personality*, 1953, 22, 142-150.
4. Kelly, E. L., "Marital Compatibility as Related to Personality Traits of Husbands and Wives as Rated by Self and Spouse," *Journal of Social Psychology*, 1941, 13, 193-198.
5. LaForge, R., and R. Suczek, "An Interpersonal Check List," *Journal of Personality*, 1955, 24, 94-112.
6. Leary, T., *Multilevel Measurement of Interpersonal Behavior*, Kaiser Foundation of Psychology Research, Berkeley, California: Psychological Consultation Service, 1956.
7. Locke, H. J., *Predicting Adjustment in Marriage*, New York: Holt and Company, 1951.
8. Locke, H. J., and K. M. Wallace, "Short Marital-Adjustment and Prediction Tests: Their Reliability and Validity," *Marriage and Family Living*, 1959, 21, 251-255.
9. Luckey, E. B., "Implications for Marriage Counseling of Self-Perceptions and Spouse Perceptions," *Journal of Counseling Psychology*, 1960, 7, 3-9.
10. Luckey, E. B., "Marital Satisfaction and Congruent Self-Spouse Concepts," *Social Forces*, 1960, 39, 153-157.
11. Luckey, E. B., "Marital Satisfaction and Parent Concepts," *Journal of Consulting Psychology*, 1960, 24, 195-204.
12. Mangus, A. R., "Integration of Theory, Research, and Family Counseling Practice," *Marriage and Family Living*, 1957, 19, 81-88.
13. Mangus, A. R., "Role Theory and Marriage Counseling," *Social Forces*, 1957, 35, 200-209.
14. Sears, R. R., "A Theoretical Framework for Personality and Social Behavior," *American Psychologist*, 1951, 9, 476-483.
15. Terman, L. M., *Psychological Factors in Marital Happiness*, New York: McGraw-Hill, 1938.

## Class Origin of Scientists<sup>1</sup>

S. STEWART WEST, *University of California at Los Angeles*

Persons who produce new knowledge through scientific research have many characteristics in common, some of which are related to the requirements of the activity itself and some to the social matrix in which it is embedded. In reaching their occupation, they undergo a complex process of selection, determined by inherent ability, by characteristics of personality resulting from early socialization, by resources of money or opportunity to buy training with it, and by accidental placement in time or space. Parental class is related to each of these factors. Nevertheless, it is not very clear what obstacles oppose upward movement in this case, at what points in a scientific career they are effective, nor how they are related to experience in the parental environment.

For a substantial fraction of individuals who will enter research, occupational plans and level of aspiration, in the sense of desire for prestige or money, are almost irrelevant. Motivated by need for cognition (need for understanding, or curiosity), these persons learn for the simple enjoyment of learning, and at each career choice-point they tend to select the alternative which makes further learning possible, until a limit of resources or ability is reached. Even in the case of those who are attracted by the prestige of research and by the possibility of a fairly high income earned in a pleasant situation, there is reason to believe that need for cognition dominates the motivation of nearly all who actually enter the occupation. One would expect such a person to pursue education to as high a level as possible and then seek a situation in which he can himself produce new knowledge, because at every step of his progress need for cognition is being satisfied.

It is by no means certain that average ability to learn and apply the cognitive skills used in research varies from class to class. Not only can one not yet specify precisely what skills are needed, but measures of intelligence have been found to be biased in favor of upper-class persons. Eells *et al.* (5) found that test items showing large status differences were those involving verbal symbolism, largely associated with a relatively academic or bookish vocabulary, or with words, concepts, or objects outside the experiences of lower-class children. It is not yet known how much of the variation of measured intelligence with social class can be accounted for by such bias (10). Hence one cannot safely assert that inherent ability to *manipulate* concepts occurs less frequently

<sup>1</sup> Data for this paper were collected in collaboration with Dr. Donald C. Pelz of the Institute for Social Research, University of Michigan. The work was supported by a grant from the Carnegie Corporation of New York.

among lower-class than among upper-class children. The quite strong relation between educational achievement and parental class which Terman and Oden (22) found in their sample of gifted men indicates also that such ability does not guarantee achievement of professional status in the same degree for all classes.

The study reported here examines effects of parental class in determining who will become research scientists,<sup>2</sup> using a sample of persons employed in research. Its intent was two-fold: to discover what characteristics of personnel differentiate research organizations, and to identify mechanisms which govern the selection of individuals for training in science. A retrospective design was used because the phenomena of selection were of interest chiefly as they determine the characteristics of men who are ultimately employed by research organizations. One wishes to know primarily what kind of person produces new knowledge, in order to predict the nature and quantity of the product.

#### SAMPLE

The data were obtained as part of a larger study of research organizations by means of questionnaires administered to the entire staff of each of five industrial laboratories and seven science departments of a large midwestern university, including all persons classified in scientific grades by the industrial organizations and all persons of faculty rank in the university departments. Numbers of questionnaires and characteristics of organizations are summarized in Table 1.

TABLE 1  
*Composition of Sample*

Research Organization	I	II	III	IV	V	VI	Total
Type of research *	R	R,A,D	A,D	A,D	D	D	—
% persons with doctorate	97.7	43.4	11.8	3.7	2.5	0	34.7
Total persons in research (total questionnaires)	238	218	60	213	129	70	928
Completed questionnaires with father's occupation	176	201	51	180	110	59	777
No. in mathematics, physics, or chemistry	79	112	51	180	110	59	591
No. in biological science	38	89	0	0	0	0	127
No. in social science	59	0	0	0	0	0	59

\* R denotes basic research. A denotes applied research. D denotes development.

<sup>2</sup> The term "scientist" is used to mean any person engaged in the production of new knowledge, either in basic research or in applied research or development, including research engineers.

TABLE 2  
*Distribution of Sample in Father's Occupational Class*

Category	Research Organization	Number of Persons with Fathers in Stated Class								
		2	3	4	F	5	6	7	8 or 9	Total
With Doctorate	I	0	7	16	11	57	41	39	1	172
	II	4	4	8	9	23	23	14	2	87
	III	—	—	—	2	1	2	1	—	6
	IV	—	—	1	1	1	1	1	2	7
	V	—	—	—	—	2	1	—	—	3
	VI	—	—	—	—	—	—	—	—	0
	Total	4	11	25	23	84	68	55	5	275
Without Doctorate	I	—	—	—	—	2	2	—	—	4
	II	1	15	23	17	14	25	17	2	114
	III	2	4	10	6	9	9	5	0	45
	IV	5	25	43	5	55	24	15	1	173
	V	2	14	19	7	31	23	10	1	107
	VI	—	6	11	4	23	8	7	—	59
	Total	10	64	106	39	134	91	54	4	502
Some graduate training (82% to master's degree but not beyond)										
		1	15	25	14	26	17	18	1	117
B.A. or B.S.										
		9	47	71	17	94	66	34	3	341
No college degree (75% with high school only)										
		0	2	10	8	14	8	2	0	44
Total persons in sample										
		14	75	131	62	218	159	109	9	777
% of sample										
		1.8	9.7	16.9	8.0	28.1	20.5	14.0	1.2	100.0
% of persons with doctorate										
		—	15.0	19.0	37.0	39.0	43.0	50.0	—	—

Father's occupation (at the time the respondent was in college) was classified according to the scale described by Kinsey, Pomeroy, and Martin (11, pp. 77-79), in part to obtain direct comparability with their data on the social mobility of a sample of 2945 males from the same general region as that from which our sample was drawn (11, p. 418). Distributions of respondents in father's class are shown in Table 2 for the persons with the doctorate and persons without, in each of the six organizations. Class 2 contains unskilled labor, class 3 semi-skilled labor, class 4 skilled labor, class 5 lower white-collar occupations, class 6 upper white-collar occupations, class 7 professionals, class 8 major business executives, and class 9 the independently wealthy. Farmers, who normally fall in classes 4 and 5, are here distinguished as a separate class.

The sample is strictly representative only of the cooperating research organizations. These lie in the northeastern quarter of the United States and have

sizes reasonably characteristic of those industrial organizations whose scientific employees are used for research and for fundamental improvement of products rather than for optimization and control of manufacturing processes. Departments surveyed at the university (Organization I) were anatomy, biochemistry, mathematics, physics, physiology, psychology, and sociology—all in basic science.

It is doubtful that a genuine probability sample could be drawn on the basis of present knowledge of relevant variables, and a survey of the total research population, even if feasible, would be hard to justify. The similarity of our six sub-samples suggests that variation within an organization greatly exceeds variation between organizations. Since inferences are chiefly to be drawn from intra-sample comparison, the sample may therefore be regarded as adequate for the purpose in view.

#### DISTRIBUTIONS IN FATHER'S OCCUPATIONAL CLASS

##### *Differences between Organizations*

Comparison of the distributions in Table 2 indicates that research organizations differ much more in the amount of training they require of their employees than in respect to the distribution of these researchers in class of origin. Chi-square tests show no significant differences between any two of Organizations II to VI in respect to the distribution of persons without the doctorate, and no significant difference between Organization I (the university) and the total of Organizations II to VI (industrial laboratories) in respect to distribution of persons with the doctorate. Here class 2 was combined with class 3, and classes 8 and 9 with class 7, leaving five degrees of freedom. The chi-square probability lay between 0.3 and 0.7 for all the above comparisons. It is therefore appropriate to combine distributions of persons with the doctorate and combine distributions of persons without the doctorate. The combined distributions differ at the .001 level of chi-square.

This is to say (on the basis of the data of Table 2) that the university is indistinguishable from the industrial laboratories and the industrial laboratories from one another, when one controls for differing proportion of persons with the doctorate. However, probability of possessing a doctorate appears to be related to parental class, as may be shown by computing the fraction of persons from each class who have the doctorate (last line of Table 2). This fraction shows a sharp discontinuity at the boundary between manual and white-collar workers, and increases only gradually from that boundary toward higher classes. It is, however, almost as large for class F (farmers) as for the white-collar classes, so that in this respect the farming fathers resemble persons of class 5 rather than class 4. This is reasonable in that most responses suggest farms which represent rather large capital investments.

*Probability of Achieving the Doctorate*

The data do not permit finer discrimination than that between manual-labor and white-collar classes; that is, between a stratum A containing classes 2, 3, and 4, and a stratum B containing classes F, 5, 6, 7, 8, and 9. Within each stratum, the distribution in class of persons with the doctorate is indistinguishable by chi-square test from that of persons without, the chi-square probability being 0.43 (at 2 df) for stratum A, and 0.64 (at 4 df) for stratum B. Of 220 persons in stratum A, 40 had the doctorate, while of 557 persons in stratum B, 235 had the doctorate, or 18.2 per cent and 42.2 per cent, respectively.

Delay in completion of undergraduate study suggests itself as an explanation of the markedly lower probability of possessing a doctorate among children of manual workers. In our sample, probability of a doctorate decreases exponentially with age at bachelor's degree, by almost tenfold per decade (Table 3). The ratio of number of persons with the doctorate to total persons who received a bachelor's degree at a given age of  $t$  years can be expressed as  $e^{-(t-18)/4.6}$ , which fits the observations very well, with chi-square probability

TABLE 3  
*Distribution of Sample in Age at Bachelor's Degree*

Age $t$ at B.A. or B.S.	Number of Persons				Probability of Doctorate (from Fitted Curve)
	Total (n)	With Doctorate			
		Observed	Predicted	Diff.	
18	1	1	1.0	0.0	1.00
19	4	4	3.2	0.8	.80
20	33	20	21.3	-1.3	.65
21	93	46	48.6	-2.3	.52
22	185	80	76.9	3.1	.42
23	96	35	32.3	2.7	.34
24	66	18	17.8	0.2	.27
25	67	15	14.5	0.5	.22
26	53	4	9.3	-5.3	.18
27	24	5	3.4	1.6	.14
28	22	3	2.5	0.5	.11
29	20	1	1.8	-0.8	.09
30	12	1	0.9	0.1	.07
31	10	1	0.6	0.4	.06
32-42	13	0	1.0	-1.0	.05 to .005
Over 42	0	0	0	0	—
No B. S.	44	0	0	0	—
Total	743	252	—	—	—

Predicted number with doctorate =  $n e^{-(t-18)/4.6}$  by fitted curve. Totals are here reduced by missing data.



larger than 0.9. This ratio is unity at 18 years of age, and decreases to .074 at age 30. Some error is introduced by neglect of persons who will later obtain doctorates by leaving their jobs to return to school or by part-time study, but this error may be expected to be small.

Children of manual workers do graduate from college somewhat later than children of white-collar workers and farmers, as shown by the distributions in Table 4. However, it appears also that the manual/white-collar differential in fraction having the doctorate has little or no dependence upon age at receipt of bachelor's degree, since expected values of number of persons with the

TABLE 4  
*Effect of Age at Bachelor's Degree on Relation of  
Probability of Doctorate to Parental Class*

Age at B.A. or B.S.	Number of Persons in Class F or 5 to 9		Number of Persons in Class 2, 3, or 4			
	(a) Total	(b) With Doctorate	(c) Total	With Doctorate		
				(d) Observed	Expected	Diff.
18 to 20	32	23	5	2	1.8	0.2
21	73	41	20	5	5.6	-0.6
22	138	69	43	9	10.8	-1.8
23	59	25	33	8	7.0	1.0
24	42	14	22	4	3.7	0.3
25	43	11	19	3	2.4	0.6
26 or more	78	10	61	3	3.9	-0.9
Total	465	193	203	34	35.2	—

The predicted number of doctorates among persons from classes 2 to 4 is computed on the assumption that the probability of a doctorate is half the probability among persons from classes F to 9. Totals are reduced by missing data.

doctorate computed according to a hypothesis of independence do not differ significantly from the observed numbers. This is a stronger statement than to say that, when age at bachelor's degree is controlled, the overall class-differentiation does not change, and it is worth demonstrating in detail.

Referring to the lettered columns of Table 4, let  $a$  be the total number of persons and  $b$  the number with the doctorate in stratum B in a given interval of age at receipt of bachelor's degree, and let  $c$  and  $d$  be the corresponding numbers of persons in stratum A. The hypothesis that the ratio of the probability of possessing a doctorate in stratum A to the corresponding probability in stratum B is independent of age at receipt of the bachelor's degree then requires that  $d/c = k(b/a)$ , or  $d = kbc/a$ , where  $k$  has the same value for any age.

Expected values of  $d$  computed from this formula with  $k = 1/2$  are shown in Table 4. Since comparison with observed values yields a chi-square proba-

bility of 0.74, it cannot be asserted that delay in graduation from college produces the manual/white-collar differential in probability of possessing the doctorate. Hence the data indicate that probability of possessing the doctorate, having graduated from college and chosen research as an occupation, is expressible as the product of two independent factors, one which is a function only of father's class and one which is a function only of age at graduation. Their independence seems as important as their functional forms.

#### *Persons without the Doctorate*

Division of the non-doctorate category of respondents into persons without college degree, persons with B.A. or B.S. only, and persons with some graduate training (Table 2) indicates that parental class affects chiefly the latter portion of doctoral training. If class F (farmers) is excluded, these three component distributions in parental class do not differ significantly among themselves, although their sum does differ from the distribution of persons with the doctorate, as pointed out previously. However, children of farmers constitute 12 per cent of the high-school graduates and 18 per cent of persons with some graduate training, but only 5 per cent of those who terminated their education at the bachelor's degree. The absolute differences in number are too small to warrant more speculation than that rural children may possibly be less inclined than urban children to persevere in education merely to reach an approved terminal level. The important point here is that, for persons oriented toward research as an occupation, parental class appears to exert little selective effect between the end of high school and the mid-point of graduate training.

The 44 persons who were engaged in research without college degrees (Table 2), three-fourths of them with no formal training beyond high school, were following a quite different career sequence. They were not merely technical assistants, but were employed at the same levels of task and salary as many of the persons with college degrees. It should not be forgotten that any of the formal stages of training can be by-passed under special circumstances. One can enter college without graduating from high school, or enter graduate school without having been an undergraduate, although such events are rare. Nevertheless, it is surprising to find that the self-taught scientist has so large a place in modern research, although Visser's sample contained many such who attained eminence at an earlier period (25, p. 358-360).

#### MOTIVATION TOWARD RESEARCH

##### *Data Concerning Motivation*

Some information with regard to motivation for engaging in research was obtained in interviews conducted prior to the distribution of the question-

naires. The sample in this case consisted of approximately one-third of each of six departments of the university (Organization I, psychology omitted) and approximately half of each of Organizations II and III. In the first case the population was dichotomized about median age and about median age at receipt of doctorate, and in the latter two cases about median age and about

TABLE 5  
*List of Reasons for Engaging in Research*

Group	Inferred Motive	Order in List Presented	Statement Presented to Respondent
A	$n$ Cognition	5	I enjoy mastering a difficult problem.
		9	I want to gain a more complete understanding of my field.
		2	I enjoy thinking systematically.
B	$n$ Nurturance	10	I want to make a contribution to the general body of scientific knowledge.
		14	I want to produce results that are needed by other people.
		11	I want to make a contribution to the welfare of society.
C	$n$ Achievement	8	I want to raise my status in the company.
		13	I want to think well of myself.
		12	I want other people to think well of me.
D	$n$ Distinction	3	I want to make the universe look orderly to me.
		1	I want to know things that most other people do not know.
		6	I enjoy upsetting inadequate theories of previous investigators.
		7	I want to earn more money.
		4	I enjoy working with laboratory equipment and data.

median salary, a proportionate sub-sample being drawn randomly from each of the four cells with random replacement of the few refusals. The total sample contained 178 persons, for whom father's occupation was known from the questionnaires.

The method adopted was that of direct self-ratings, obtained by presenting to the respondent a list of reasons for engaging in research (Table 5), from which he was asked to select the three most typical of his own reasons, and then the three least typical. This list was based upon reasons offered in open-

ended interviews of a previous study, extended by addition of implied alternatives.

Although there are difficulties in interpreting such data, as McClelland has pointed out (1, pp. 7-42), more refined and indirect tests appear not to be available for all motives which might be relevant to research. Previous investigators have devised methods for measuring need for achievement, need for affiliation, and need for power by means of fantasy-response patterns (1), and at least one attempt has been made to measure need for cognition (3), but other motives must also be considered. Moreover, protocols presented by Roe (16, 17) from the pictures of the Thematic Apperception Test suggest that to scientists these pictures have an alien quality which inhibits response. In the present state of technique, only the method of direct report seemed sufficiently general for the purpose at hand.

Since one wishes to make inferences to underlying motives, it is useful to group consonant reasons at a level of abstraction comparable to that of motives discussed in the literature. The grouping was considered to be optimum when contradictions in mention had been minimized, where *contradiction* means that a respondent has cited one statement of a group as most typical of himself and another as least typical.

In the final arrangement defined by the lettered groups, perceived dissonance of the statements (as indicated by contradictions) is small: For approximately one person in 20 who mentioned statement 2, 5, or 9, wanting to gain a more complete understanding of one's field was not consonant with thinking systematically or with mastering difficult problems (Group A). For approximately one person in 20 who mentioned 10, 11, or 12, contributing to the general body of scientific knowledge, contributing to the welfare of society, and producing results needed by other people were not mutually consonant (Group B). For approximately one person in ten who mentioned statement 8, raising one's status was not consonant with wanting other people to think well of one or with wanting to think well of oneself (Group C). For approximately one person in 20 who mentioned statement 1, 3, or 6, making the universe look orderly was not consonant with wanting to know things others do not know or with upsetting inadequate theories (Group D).

Statement 7 ("I want to earn more money") was in contradiction with statements of Group C and Group D a third of the times it was cited, respondents apparently perceiving it as quite different from either of these two groups. Statement 4 ("I enjoy working with laboratory equipment and data") was described as non-applicable by many persons (such as mathematicians or sociologists) whose disciplines or problem areas did not include laboratory situations, so responses to it were excluded as not being comparable across disciplines.

*Designation of Motives*

The level of abstraction represented by the lettered groups of Table 5 approximates that of motives described in the literature. Group A corresponds quite closely to the *need for cognition* studied by Cohen, Stotland, and Wolfe (3), or Murray's *need for understanding* (8, 15), and is similar to the exploration (or curiosity) drive which was identified by Cattell (2, p. 516). Group B is essentially the science-relevant aspect of Murray's *need for nurturance*, and resembles what Cattell has called parental protectiveness (2, pp. 515-516). Group C may be identified with *need for achievement*, since its statements imply competition against some standard of excellence (1, p. 181). Statements 12 and 13, however, are more general than statement 8, for one might desire self-approval or group approval not mediated by such an indicator of success as status.

Since the motives whose presence is suggested by the lettered groups of statements are so nearly the same as those which have been given names by other investigators, they will be referred to by these names: Group A as *n* Cognition, Group B as *n* Nurturance, and Group C as *n* Achievement. The agreement, if not complete, is very close.

Each of these three needs appears to include a component of striving to modify the environment. In contrast, the statements of Group D suggest an ideational restructuring, by ordering one's perceptions, by obtaining rare knowledge merely for the pleasure of possessing it, or by demonstrating the falsity of other men's theories. Here the focus is on a passive end-state, rather than on answering questions or solving problems—a contrast of scholar with researcher. For want of an accepted designation, the motive implied by Group D will be referred to as *need for distinction*.

*Relative Frequencies of Motives*

Table 6 shows numbers of persons who cited statements without contradiction in each of the four groups. Motives inferred as above from reports of the respondents are ranked in the following decreasing order of importance: *n* Cognition, *n* Nurturance, *n* Achievement, and *n* Distinction. Seventy-eight per cent of respondents asserted *n* Cognition and 81 per cent denied *n* Distinction.

Questionnaire data obtained by Visser (25) from 825 scientists starred in *American Men of Science* show the same order of importance of motives, insofar as comparison is possible. Closest approximation of the conceptual frameworks of the two studies is obtained under the following assumptions: (a) that Visser's categories of great, moderate, or slight stimulation by the motive in question correspond to our response categories of most typical, non-

TABLE 6  
*Responses to Groups of Reasons for Engaging in Research*

Groups of Reasons (See Table 5)	Most Typical of Self		Non-Mention or Contradiction		Least Typical of Self	
	No.	%	No.	%	No.	%
A	138	78	29	16	11	6
B	114	64	34	19	30	17
C	35	20	71	40	72	40
D	10	6	24	13	144	81
7	18	10	99	56	61	34

mention, or least typical, and (b) that "curiosity" in Visser's scheme corresponds to Group A, "helping mankind" to Group C, "obtaining recognition" to Group C, and "rivalry" to Group D. On these assumptions, the rank-ordering of motives is the same in Visser's sample as in our sample (whose members were born about 40 years later and were not selected for eminence) and the frequencies are not very different.

Now the chief purpose for introducing these data is to discover whether there is a significant difference in motivation toward research between Stratum A (classes 2, 3, and 4) and Stratum B (classes F and 5 to 9). There appears to be none. If one divides the respondents of Table 6 into the 47 of Stratum A and the 131 of Stratum B and compares responses by  $3 \times 2$  contingency tables, each of the five motives yields a chi-square probability between 0.5 and 0.8.

However, class 7 (professionals) shows greater incidence of *n* Nurturance (Group B) and less incidence of *n* Distinction (Group D) than other classes, both significant at the .01 level of chi-square. By greater incidence is meant that the fraction of "most typical" responses is larger and the fraction of "least typical" responses is smaller. In addition, desire for money (statement 7) is cited as least typical twice as frequently by persons of class 7 as by those of other classes, a difference which is significant at the .001 level. There are no other significant differences between classes, although class 5 (lower white-collar) tends in a direction opposite to that of class 7 in each of the three motives mentioned, and persons from class F (farmers) report *n* Cognition somewhat more frequently than do those from other classes.

Differences between research organizations are of such a nature as not to be explicable on the basis of a contrast between basic and applied research. Organization I, the university, differs from Organization III, which is devoted to applied research in glass manufacture, only in a smaller incidence of desire for money (.01 level). However, in Organization II, which conducts



basic and applied research in pharmaceuticals, twice as many persons as in either I or III cited *n* Nurturance (Group B) as most typical, and half again as many as in I or III cited *n* Distinction (Group D) as least typical (both significant at the .001 level). Apparently the use to which the research is to be put has far more selective influence than the immediacy of its application or the degree to which it is channelized toward a goal.

In Organization II, the 22 supervisors were almost identical in response with the 70 non-supervisors and the 45 persons with the doctorate almost identical with the 47 persons without. Moreover, there were no significant differences among (a) the 33 only-children in the total sample, (b) the 52 first-born with sibs, (c) the 51 non-first born with one or two sibs, and (d) the 42 non-first born with three or more sibs, although both *n* Nurturance and *n* Distinction tend to increase in frequency of assertion in this order. Hence, organizational status and level of training appear to be unrelated to motivation, and birth order or family size no more than weakly related.

At the university (I), the 29 persons who had published 6 or more papers in the previous five years differed from the 24 persons who had published 5 or less papers only in more frequent denial of desire for money (.02 level). This leaves one little evidence to support a relation between motivation and ability.

#### *Interpretation of Motives for Engaging in Research*

Even when one allows for a leveling effect of group norms on report of motivation, the homogeneity of mobile persons in respect to reasons for engaging in research is rather extraordinary. Only the non-mobile class-7 persons appear to be differently motivated, and in a manner to be expected from the greater security of having been socialized within the professional class. Moreover, response to the statements of Group C indicates that *n* Achievement is little more than randomly associated with choice of research as an occupation, and even then negatively.

How the other motives (primarily *n* Cognition) may initiate and maintain upward mobility is suggested by remarks of Kinsey, Pomeroy and Martin concerning data of a quite different kind (11, pp. 418-419). They found that in general the sexual history of an individual accords with the pattern of the terminal class rather than that of the class of origin. That is, in nearly all cases the mobile individual had broken with the patterns of the parental class while still living with his parents, and then gravitated to a class whose behavior was like his own. The far rarer case in their data was that where behavior of the terminal class is learned in maturity, through a process of acculturation like that so vividly described by Shuetz (19). By analogy, it is probable that for most researchers (and for all in some degree) aspiration

has meant desire for association with persons having interests, values, and especially modes of thinking like those toward which early socialization had inclined the researchers themselves, and for situations in which these interests and values could be implemented. This is the kind of effect which one would predict also from Festinger's theory of cognitive dissonance (7).

Since research activity consists of solving problems for which data or methods or both are lacking, motivation for this activity is most probably derived from the basic psychological phenomenon of anxiety in the presence of an unstructured situation. Such anxiety can be reduced by acquiring information, by developing skills to acquire it more efficiently, by organizing it in solution of a problem, and by finding new problems to which these skills may be applied. Alternatively, it can be reduced by internalizing cultural prescriptions for behavior and thought, or by dependence on authority. In the former case, *n* Cognition or curiosity can be satisfied in such a situation as a research organization or, rarely nowadays, in independent research. In the latter case, new knowledge is not required, and so is not sought.

Habituation in obtaining new knowledge for oneself appears to imply some degree of isolation during early childhood, as was suggested by Faris (6), and reported by the science-talent contest winners studied by McCurdy (14). Roe's data support this formulation, and indicate also that parents of eminent scientists tend to be affectively cold. In the extreme, one would expect the resultant personality of such a child to be characterized by (a) impoverishment of affect, (b) lack of indoctrination in cultural concepts, (c) functional organization of concepts in thinking, (d) narrowing of interest and attention, and (e) use of fantasy to solve problems and provide satisfactions. Given some degree of social acceptance, all these characteristics contribute to effectiveness in research.

Acceptance, however, must occur in an almost continuous sequence of situations from childhood to the end of productive life, of which each will have become accessible largely by chance. If support is absent for too long a period, motivation may be destroyed. The results of an unsupported childhood have been described by Dunham (4) in a paper on the social personality of the catatonic schizophrenic. Here one should note especially the strong interest of his subjects in reading advanced books, which did them little good in the absence of opportunity to apply information thus obtained. Mature research workers can also be affected by withdrawal of support, to the physical or psychological detriment of the researcher or at least to the detriment of science.

All the conditions for the development of a research-oriented personality which have been stated or implied above become more probable with higher parental class: affectively cold parents; isolation from the physical world, from siblings, from peers, and from adults; favorable attitudes toward educa-

tion; approval of abstract thought; accessibility of congenial activity. Evidence is scattered and needs to be augmented and systematized, but the summary in Eells *et al.* (5, pp. 39-47) indicates its direction: one may expect lower-class children to develop abstract, objective, and autonomous patterns of thought less often than upper-class children, and to be more often frustrated and penalized in exercising them. The result of these phenomena would be an increase in scientific ability with increasing level of parental class.

Insofar as the Stanford-Binet intelligence test is an indicator of such patterns of thought, the data of Terman and Oden (22) provide evidence that persons possessing them tend to seek research activity. Although only about 0.2 per cent of males in the general population engage in research, 2.6 per cent of males with  $IQ \geq 135$  in this sample were listed in *American Men of Science*, and almost 15 per cent of the males with  $IQ \geq 170$  were at least employed in situations where research is customary. However, the probability of achieving such a situation was smaller for children from lower parental classes, even though mean IQ was nearly the same. Probability of doctorate in the general population has been found by Harmon (9) to increase with intelligence-test score in a similar manner, but this is not a direct indicator either of participation or ability in research.

Comparison with samples of other investigators (Table 7) indicates only a small overall effect of parental class on achievement of eminence, although a significant one as between Visser's sample and the persons with doctorate in Table 2. However, the markedly smaller contribution of classes 2 and 3 to the samples of eminent scientists raises again the question of whether class-

TABLE 7  
*Comparison with Samples of Other Investigators*

Composition of Sample *	% Fathers in Stated Class			Sample Size
	2 or 3	4, F, or 5	6 to 9	
Men with I.Q. over 135 (Terman and Oden)	8	26	66	766
Professional men (Kinsey <i>et al.</i> )	8	36	56	671
Eminent physicists and biologists (Roe)	0	42	58	42
Eminent scientists (Visser)	1	48	51	849
Researchers with doctorate, Table 2	5	48	47	275
Physicians, 1940 (Rogoff)	12	46	42	105
Researchers with some graduate training, Table 2	14	55	31	117
Researchers with four years of college, Table 2	16	54	30	341
Engineers and scientists, 1940 (Rogoff)	9	63	28	178
Researchers without college degree, Table 2	5	73	23	44
Contemporaries of persons in Table 2 (adjusted for the effect of siblings)	44	46	10	Total USA

\* For sources, see references 10, 15, 16, 17, 20, 23.

related traits of personality may not restrict the range of opportunity accessible to children of unskilled and semi-skilled fathers. Average parental class trends upward in Table 7 with amount of formal training, but it has been shown previously that by far the most critical point of selection occurs at the level of the master's degree, and that here the critical distinction in parental class is between blue-collar and white-collar fathers.

From the above discussion one may construct an interpretation of the data in the following sequence of hypotheses: (a) Motivation toward research consists predominantly of  $n$  Cognition or curiosity. (b) This is developed during childhood by some degree of social isolation. (c) It is equivalent to (or highly correlated with) scientific ability. (d) Probabilities that  $n$  Cognition will be induced, reinforced, and implemented all increase with increasing level of father's occupational class. (e) Hence probability of seeking problem-solving skills and activities, the research occupation itself, and association with research-oriented persons increases with increasing level of father's class. (f) For persons who will be employed in research, terminal level of education does not depend on intensity of  $n$  Cognition, but does depend on the attitudes of authority figures toward other attributes of personality.

The first of these statements may appear self-evident, and all but the last quite reasonable or even elementary. Yet in each case it is not so much a test of the hypothesis which is required as demonstration of a causal mechanism. This is a difficult task, in view of the large effects of random placement of persons in space and time. Hence, if the process outlined appears to be an oversimplification of conclusions reported in the literature (as reviewed, for example, by Super and Bachrach (21)), this may be advantageous in defining the core of the process of occupational selection.

#### SELECTION FOR TRAINING

##### *Selection for College*

For the sample described by Table 2, the larger part of the selective effect of parental class occurred before graduation from high school, as may be seen by comparison with a rough estimate of the distribution in class of those contemporaries of the persons in the present sample who might have become scientists (last line of Table 7). This estimate was obtained by averaging the compositions of the total male labor force of the USA in 1940 and 1950, as given by the censuses of those years (23, 24), and multiplying by an effective mean number of children adjusted for number of siblings according to computations in a previous paper (26). (It was shown in the previous paper that the probability of a given child's becoming a scientist decreases rapidly with increasing number of siblings, with the result that the mean sibship size effective for contributing scientists is smaller than the actual mean sibship size

in the general population.) The decade 1940-1950 was taken as base because the median age of respondents was 35 years in 1959 and father's occupation was specified for the period when the respondent was in college.

The foregoing section has suggested that a large part of this selection for scientific training in college can be ascribed to class-related variation of conditions which induce and reinforce motivation toward research. This is illustrated also by the finding of Kohn (13) that the fraction of mothers who value curiosity in their children increases rapidly with increasing class. It is much less likely that aspiration to college education in general depends on parental class. Wilson (27) has asserted that influence of the peer-group on level of aspiration is demonstrated by the fact that segregation in respect to parental class (by districting of high schools) has a strong effect on the fraction of students planning to go to college. According to the evidence he presents, this fraction can be twice as large for upper-class as for lower-class schools and is almost independent of academic grade and IQ score. However, Wilson's respondents were reporting *plans*, which were probably based on fairly realistic perceptions of resources. A study of ninth-graders by Stephenson (20) yielded much the same results with respect to occupational plans, but showed almost no class-related variation of occupational aspiration. This indicates that the chief determinant of planning for college in general is simply ability of parents to pay for such training. Since manual workers living in white-collar districts may be presumed to be earning more than the average for their class, their children should more often be able to plan realistically for college, as Wilson found them to do.

#### *Selection for Doctoral Training*

The one other large effect of class indicated by our sample occurs at the level of the master's degree, where the probability of going on to a doctorate in science is twice as large for children of white-collar workers or farmers as for children of manual workers. Since the income ranges of skilled-labor and clerical workers overlap, the chief determinant of the class factor in selection for doctoral training is probably not to be sought in relative ability of parents to provide funds for education. Indeed, this should be so even if parental income accurately paralleled the prestige of parental occupation, for graduate study has come to be financed chiefly by the student himself, through fellowships, assistantships, and other academic jobs. Hence one is led to consider not only the attitude of the student toward further study, but also that of professors toward the student. The attitudes of professor toward student determine not only his granting or denial of financial help in the form of fellowships or assistantships, but also whether he brings opportunities to the attention of the student or otherwise encourages him. If the professor

can identify with the student on the basis of class-related characteristics of personality, he is likely to say, "Here is a man who will become a scientist like me. He deserves to be helped." Conversely, a person who is upwardly mobile from a manual-labor family may have developed such defenses as an aggressive attitude, or carried with him a short-term perspective which demands immediate satisfaction from each learning activity. In such a case, discrimination against him can occur, regardless of his ability.

The importance of *unsought* help is indicated by the fact that 17 of Roe's 22 eminent physicist subjects specifically mentioned receiving it (16). It can be important not only in revealing unrecognized opportunities and in providing direction and encouragement, but also because initiative from above leads to a favorable outcome for the student much more often than initiative on his own part. If he is marked by unfamiliar behavior, concepts, and values, his strivings may seem unworthy of reward because of their uncouthness. Nevertheless, lower-class traits which are considered undesirable in the academic situation may be unrelated (or even positively related) to effectiveness in research. This is a matter which appears not to have been investigated.

It is, of course, a reasonable conjecture that children of manual workers should have a more utilitarian view of education than children of white-collar workers, and should tend not to defer the gratification of earning and doing during the tedious formality of a dissertation. Such an attitude would have been enhanced by the exercise of parental authority in terms of the immediate situation rather than for the development of internalized controls, on which basis Kohn found "working-class" parents to be differentiated from "middle-class" parents (12). However, among persons oriented to choose research as an occupation, and presumably motivated by need for cognition, such an attitude would still be unlikely.

#### SUMMARY

Faculty members in seven science departments of a large university and research scientists in five industrial laboratories were found to be distributed in father's occupational class in such a way that (a) there were no significant differences between organizations when the fraction of persons having the doctorate was controlled, and (b) there were no significant differences (when children of farmers were excluded) between terminal levels of education from high-school graduation to master's degree. The fraction of persons having the doctorate averaged 18 per cent for children of manual workers, and 42 per cent for children of white-collar workers and farmers. In either of these categories, there was no significant difference between the distribution of persons with the doctorate in father's class and that of persons without the doctorate.



In this sample of persons engaged in research, the fraction of persons possessing the doctorate decreased exponentially with age at bachelor's degree, from unity at 18 years to 0.07 at 30 years. This relation was independent of the effect of father's class.

Reports of motivation for engaging in research, obtained from a subsample of the respondents, showed a well-defined hierarchy of motives, but no significant differences in report of the four major motives in relation to (a) possession of the doctorate, (b) basic vs. applied research, (c) blue-collar father vs. white-collar father, (d) organizational status, (e) birth order or number of siblings, or (f) productivity among persons at the university. However, non-mobile persons (having professional fathers) did differ significantly in reported motivation from persons who had been mobile in class.

The major observations were therefore (a) that the researchers were quite homogeneous in reported motivation, (b) that father's class produced little selective effect between graduation from high school and the level of the master's degree, (c) that strong selective effects occurred below and above this period of life, and (d) that the latter selection was chiefly characterized by a contrast between blue-collar and white-collar fathers. As an interpretation of these phenomena, it is suggested: that there exist certain motives and modes of thinking characteristic of research; that these are induced during childhood (with probabilities directly related to parents' occupational class); that persons in whom they have been induced will seek the congenial activities of learning and (ultimately) employment in research; and that selection at the highest level of graduate training is produced largely by other characteristics of personality, related to parental class but not to ability in research.

*Manuscript received: October 18, 1960*

*Revised Manuscript received: April 6, 1961*

S. Stewart West  
School of Nursing  
University of California Medical Center  
Los Angeles 24, California

#### REFERENCES

1. Atkinson, J. W., (ed.), *Motives in Fantasy, Action and Society*, Princeton, N. J.: Van Nostrand, 1958.
2. Cattell, R. B., *Personality and Motivation Structure and Measurement*, Yonkers-on-Hudson, New York: World Book Co., 1957.
3. Cohen, A. R., E. Stotland, and D. M. Wolfe, "An Experimental Investigation of Need for Cognition," *Journal of Abnormal and Social Psychology*, 1955, 51, 291-294.
4. Dunham, H. W., "The Social Personality of the Catatonic Schizophrenic," *American Journal of Sociology*, 1943, 49, 508-518.

5. Eells, K., A. Davis, R. J. Havighurst, V. E. Herrick, and R. W. Tyler, *Intelligence and Cultural Differences*, Chicago: University of Chicago Press, 1951.
6. Faris, R. E. L., "Sociological Causes of Genius," *American Sociological Review*, 1940, 5, 689-699.
7. Festinger, L., *A Theory of Cognitive Dissonance*, Evanston, Ill.: Row, Peterson and Co., 1957.
8. Hall, C. S., and G. Lindzey, *Theories of Personality*, New York: Wiley, 1957. Pp. 173-174.
9. Harmon, L. R., "High School Backgrounds of Science Doctorates," *Science*, 1961, 133, 679-688.
10. Jones, H. E., "The Environment and Mental Development," in L. Carmichael (ed.), *Manual of Child Psychology*, (2nd ed.), New York: John Wiley and Sons, 1954. Chapter 10.
11. Kinsey, A. C., W. B. Pomeroy, and C. E. Martin, *Sexual Behavior in the Human Male*, Philadelphia: W. B. Saunders Co., 1948.
12. Kohn, M. L., "Social Class and the Exercise of Parental Authority," *American Sociological Review*, 1959, 24, 352-366.
13. Kohn, M. L., "Social Class and Parental Values," *American Journal of Sociology*, 1958-9, 64, 337-351.
14. McCurdy, R. D., "Characteristics and Background of Science Students," *School Review*, 1956, 64, 67-71.
15. Murray, H. A., *Exploration in Personality*, New York: Oxford, 1938.
16. Roe, A., "A Psychological Study of Eminent Biologists," *Psychological Monographs*, 1951, 65 (no. 14, whole no. 331), 1-68.
17. Roe, A., "Psychological Study of Physical Scientists," *Genetic Psychology Monographs*, 1951, 43, 121-239.
18. Rogoff, N., *Recent Trends in Occupational Mobility*, Glencoe, Ill.: The Free Press, 1953.
19. Shuetz, A., "The Stranger: An Essay in Social Psychology," *American Journal of Sociology*, 1943-4, 49, 499-507.
20. Stephenson, R. M., "Mobility Orientation and Stratification of 1000 Ninth Graders," *American Sociological Review*, 1957, 22, 204-212.
21. Super, D. E., and P. B. Bachrach, *Scientific Careers and Vocational Development Theory*, New York: Columbia University, 1957.
22. Terman, L. M., and M. H. Oden, *The Gifted Child Grows Up*, Stanford, California: Stanford University Press, 1947.
23. U. S. Dept. of Commerce, Bureau of the Census, *Sixteenth Census of the United States: 1940*, Vol. II: *Characteristics of the Population*, Part 1: *United States Summary*, Washington: U. S. Government Printing Office, 1943. Table 43, p. 95.
24. U. S. Dept. of Commerce, Bureau of the Census, *Census of Population: 1950*, Vol. II: *Characteristics of the Population*, Part 1: *United States Summary*, Washington: U. S. Government Printing Office, 1953. Table 77, p. 129.
25. Visher, S. S., *Scientists Starred 1903-1943 in American Men of Science*, Baltimore: Johns Hopkins Press, 1947.
26. West, S. S., "Sibling Configurations of Scientists," *American Journal of Sociology*, 1960-61, 66, 268-274.
27. Wilson, A. B., "Residential Segregation of Social Classes and Aspirations of High School Boys," *American Sociological Review*, 1959, 24, 836-845.

## The Perceived Structure of Political Relationships<sup>1</sup>

SAMUEL MESSICK, *Educational Testing Service*

The present report describes an initial phase of a research project which attempts to determine the extent to which perceived similarities and differences expressed in political judgments of various kinds are related to agreements and disagreements about basic social values. By "basic social value" is meant any statement defining an ethical or morally proper course of action in a situation having crucial implications for the functioning of a society. There are probably a limited set of social values in this sense which most Americans uphold in common and which limit the range of American political opinions. At the same time, however, there are also several value *differences* of a fundamental nature, which give rise to many of the cleavages in American political life. These value differences may provide stable points of reference underlying a multitude of political disputes and controversies, even those cutting across party lines, and may give continuity to political positions throughout a series of changing issues from decade to decade.

Our main study deals with such value differences and their possible relationships to political judgments and orientations. The present report, however, will focus upon the structure of perceived political relationships, as derived from judgments of similarity among 20 political leaders with respect to their political thinking. The purposes of this segment of the project were (a) to ascertain the number and nature of dimensions required to summarize the similarity judgments and (b) to investigate possible group differences in perceptual structures in relation to political affiliation.

### METHOD

#### *Subjects*

A 33-page questionnaire was administered to 836 undergraduates, 574 males and 262 females, at the University of Illinois in the Spring of 1955. The sample

<sup>1</sup> The present study is part of a larger project on relationships between political orientations and social values done in collaboration with Dr. Martin U. Martel, Department of Sociology, State University of Iowa. The research was supported in part by the Department of Psychology, University of Illinois, and in part by Educational Testing Service. The project was initiated while the author was a Ford Foundation Postdoctoral Fellow at the University of Illinois. Grateful acknowledgment is extended to Professors Lyle Lanier and J. McV. Hunt for their continuing assistance and encouragement, to Dr. C. W. Roberts for providing subjects and classroom time for data collection, to Dr. Ledyard Tucker for his invaluable advice and insights concerning the analysis, and to Dr. Carl Helm and Miss Henrietta Gallagher for supervising the computations.

consisted of 761 freshmen, 58 sophomores, and 17 upperclassmen and included representatives from colleges of agriculture (100), commerce and business administration (142), education (65), engineering (176), fine and applied arts (71), and liberal arts and sciences (250).

### *Procedure*

One section of the questionnaire contained a list of all possible pairs of 20 political leaders:

- |                                  |                                   |
|----------------------------------|-----------------------------------|
| 1. Chiang Kai-shek               | 11. Senator McCarthy of Wisconsin |
| 2. Thomas Dewey                  | 12. Prime Minister Nehru of India |
| 3. Senator Dirksen of Illinois   | 13. Richard Nixon                 |
| 4. Senator Douglas of Illinois   | 14. Franklin D. Roosevelt         |
| 5. Dwight Eisenhower             | 15. Joseph Stalin                 |
| 6. Senator George of Georgia     | 16. Adlai Stevenson               |
| 7. Alger Hiss                    | 17. Senator Taft of Ohio          |
| 8. Adolph Hitler                 | 18. Governor Talmadge of Georgia  |
| 9. Senator Kefauver of Tennessee | 19. Harry Truman                  |
| 10. General Douglas MacArthur    | 20. Henry Wallace                 |

The subjects were required to rate on a nine-point scale the similarity of the members of each pair with respect to their political thinking.

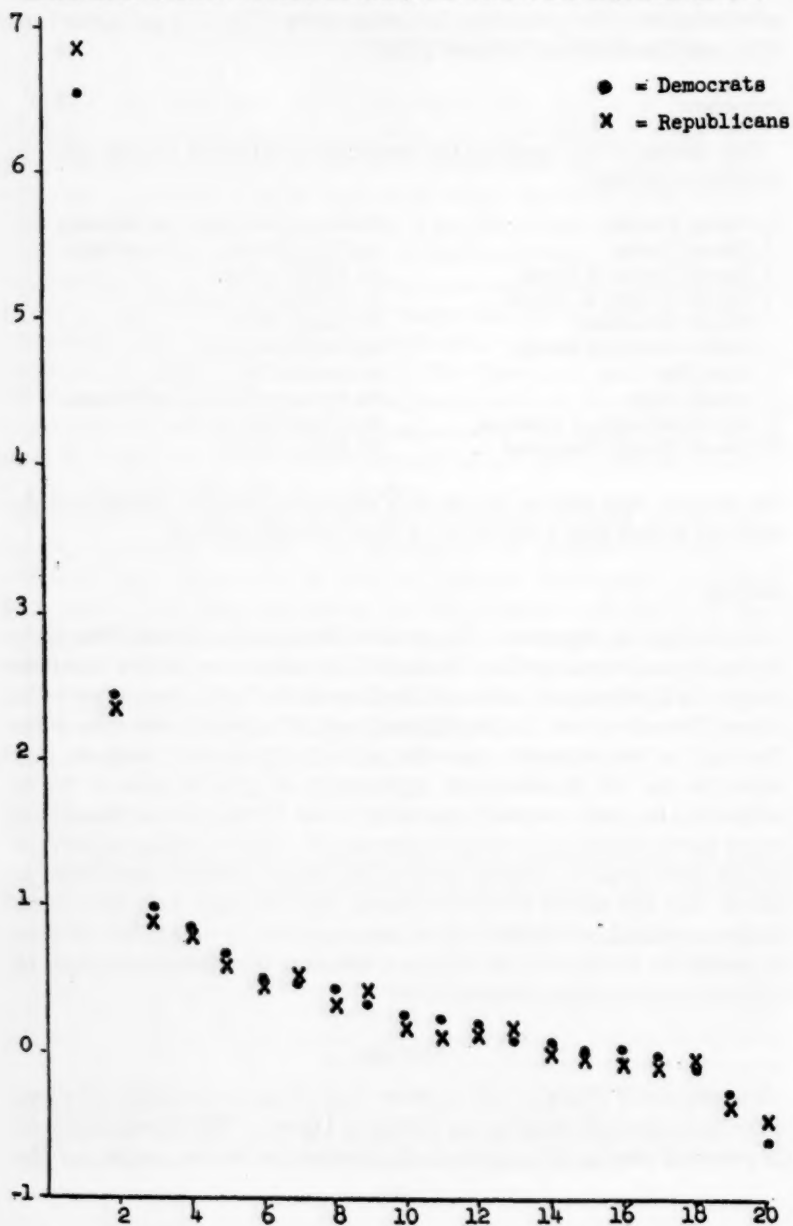
### *Analysis*

On the basis of responses to the question "In general, which political party do you support in most political matters?" the sample was divided into three groups—267 students who endorsed the Democratic Party, 464 subjects who aligned themselves with the Republicans, and 105 students who were either "not sure" or who supported some other political organization. Similarity scale values for the 190 stimulus-pairs, representing all possible pairs of the 20 political figures, were computed separately for the Democratic and Republican groups by the method of successive intervals (2), thus providing estimates of relative inter-stimulus distances between the various political leaders used as stimuli. The two sets of distance estimates were adjusted with appropriate additive constants, and scalar products between distances were factor-analyzed separately for the two groups of judges, following the general procedures of multidimensional scaling analysis (5, 6).

### RESULTS

The two sets of 20 latent roots obtained from principal-components analyses of the scalar products matrices are plotted in Figure 1. The dimensionality of the perceived relationships appears to be identical for the two parties, and the

FIGURE 1

*Latent Roots of Scalar Products Matrices for Democratic and Republican Samples*

relative sizes of the roots suggests that two dimensions would be sufficient to account for most of the inter-point distance variance. However, since stimulus loadings for some of the smaller dimensions appeared to be very similar in the Republican and Democratic solutions, it was difficult to discount these smaller vectors as random. In order to ascertain the stability of the small dimensions across the two samples of judges, ten factors from each structure were simultaneously rotated to a matching position by Tucker's (7) factor synthesis procedure. As seen in Table 1, nine of the dimensions displayed a

TABLE 1  
*Coefficients of Congruence for Matching Dimensions  
in Democratic and Republican Samples*

	$\phi^D$	$\phi^R$
1.	.99997	.99998
2.	.99719	.99859
3.	.99625	.99812
4.	.99294	.99646
5.	.98707	.99351
6.	.95985	.97972
7.	.92922	.96396
8.	.84624	.91991
9.	.71668	.84657
10.	.00510	.07141

substantial degree of similarity in the two samples, seven of the "coefficients of congruence"  $\phi$  (7), which are analogous to canonical correlations, being above .96. This coefficient of congruence is very similar to the product-moment correlation between the loadings on a particular rotated factor as it appears in the Democratic analysis and loadings on the corresponding factor in the Republican solution, after both spaces have been rotated to maximize congruence. Coefficients of congruence above .90 are interpreted as indicating a high degree of relationship. Since the matching technique involved a maximization procedure which may have capitalized upon chance relationships, a conservative estimate of the number of stable dimensions was adopted, and only seven factors were retained for rotation.

Stimulus loadings on each Democratic factor were averaged with the corresponding loadings on the matched Republican factor, and the seven combined dimensions were rotated manually by two-dimensional graphs to oblique simple structure. A small translation was also made to place the origin at the intersection of stimulus arrays. Scale values for the 20 stimuli on each of the seven rotated dimensions are given in Table 2, and the cosines between reference vectors appear in Table 3.



TABLE 2

*Scale Values for 20 Political Leaders on Seven Rotated Dimensions  
for Both Parties Combined*

	Dimension						
	I	II	III	IV	V	VI	VII
1. Chiang Kai-shek	04	-27	00	-21	-08	36	-06
2. Thomas Dewey	25	01	08	22	16	-14	16
3. Senator Dirksen	17	04	09	-09	-07	-17	29
4. Senator Douglas	-34	-06	-07	-02	12	15	15
5. Dwight Eisenhower	23	02	-13	07	01	14	-11
6. Senator George	-33	06	-05	00	02	01	36
7. Alger Hiss	-01	-27	70	-16	05	-20	-01
8. Adolph Hitler	-01	-61	46	02	-11	00	50
9. Senator Kefauver	-27	17	02	-11	-11	-10	-19
10. Douglas MacArthur	00	02	-03	01	-02	52	00
11. Senator McCarthy	14	-02	07	08	-77	06	05
12. Prime Minister Nehru	11	-49	-01	-06	00	-11	-11
13. Richard Nixon	31	18	01	-06	-16	-01	06
14. Franklin Roosevelt	-38	-06	07	-37	13	20	-12
15. Joseph Stalin	-18	-66	60	-06	01	08	04
16. Adlai Stevenson	-33	-04	-14	-22	10	-03	24
17. Senator Taft	26	-03	03	26	-15	10	02
18. Governor Talmadge	-30	-05	-07	23	-17	-04	-02
19. Harry Truman	-30	01	06	-36	-01	-21	-10
20. Henry Wallace	-20	01	63	-33	04	12	-04

On the first dimension listed in Table 2, positive scale values were obtained for Dewey, Dirksen, Eisenhower, McCarthy, Nixon, and Taft, and negative values were obtained for Douglas, George, Kefauver, Roosevelt, Stevenson, Talmadge, Truman, and Wallace. Since the Republican stimulus figures were thus cleanly separated from the Democratic stimuli and since the foreigners

TABLE 3

*Cosines Among Reference Vectors*

	I	II	III	IV	V	VI	VII
I	1.00	.08	-.16	.54	-.25	.01	-.08
II	.08	1.00	-.35	.09	-.12	-.12	-.19
III	-.16	-.35	1.00	-.49	.14	.14	.03
IV	.54	.09	-.49	1.00	-.24	-.03	.18
V	-.25	-.12	.14	-.24	1.00	.10	.10
VI	.01	-.12	.14	-.03	.10	1.00	-.16
VII	-.08	-.19	.03	.18	.10	-.16	1.00

and politically inactive figures (Hiss and MacArthur) received negligible loadings, this factor was tentatively identified as a Republican vs. Democrat dimension.

On dimension II in Table 2, high loadings were obtained for Stalin, Hitler, and Nehru, and moderate loadings for Chiang and Hiss, thus separating all of the foreign stimulus figures, along with Hiss, from the American politicians. Accordingly, this dimension of perceived stimulus difference was tentatively interpreted as a Foreigner factor or, more broadly, as a Foreign Ideology factor.

On dimension III, very high values were obtained for Hiss, Wallace, Stalin, and Hitler, suggesting a factor primarily evaluative in nature, perhaps with an emphasis upon safe vs. dangerous qualities.

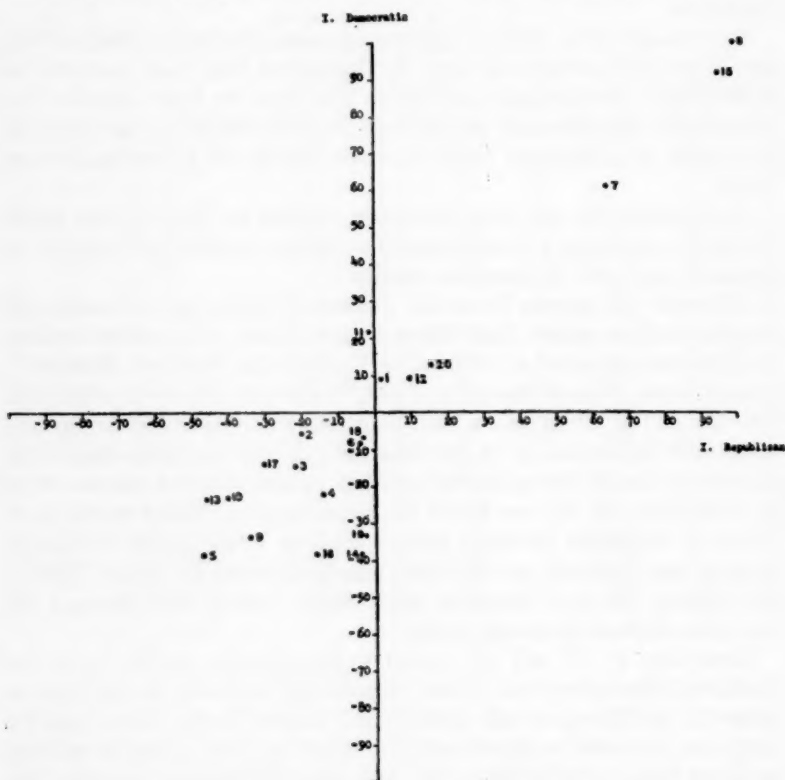
Dimension IV opposes Roosevelt, Truman, Wallace, and Stevenson with negative loadings against Taft, Talmadge, and Dewey with positive loadings, an alignment suggestive of a Liberal vs. Conservative dimension. However, a rotation in the plane of dimensions I and IV offers an alternative orientation. The plot of I vs. IV reveals a fairly unsatisfactory simple structure and an undesirably high cosine of .54, but since the rival rotation mildly disrupts the simple structure of other plots, both positions can be supported and are offered as alternatives. On the new factor IV, a high positive loading would be retained for Talmadge, moderate positive loadings would appear for George, Douglas, and Kefauver, and moderate negative loadings for Nixon, Dirksen, and Chiang. This new dimension might reflect, among other things, a not very clean Southern Democrat factor.

Dimensions V, VI, and VII appear to be primarily specific factors for McCarthy, MacArthur, and Hitler, respectively. However, the loadings on factor VI for Chiang in one direction and Truman in the other suggest a dimension somewhat broader than a MacArthur specific—perhaps a Military or a Far Eastern Policy dimension. Also, the appearance of three or four moderate loadings along with Hitler on factor VII in a combination that offers no compelling interpretation suggests that the possible inclusion of one or two additional factors in the rotation might have clarified this factor.

#### DISCUSSION

The present procedure of averaging stimulus loadings on corresponding dimensions from the Republican and Democratic samples of course produces a combined structure that reflects similar properties perceived in common by the two groups of judges. In view of the size of coefficients of congruence between matching factors in the two samples, however, there was little choice but to emphasize the common perceived structure; any differences emerging from a comparison of separate Republican and Democratic spaces would be

FIGURE 2

*Comparison of Largest Dimension from Republican and Democratic Samples*

very subtle indeed with such high canonical coefficients. To illustrate this, the first five dimensions were independently rotated for each party to a reasonably comparable orthogonal position by the quartimax procedure. The largest Democratic dimension from this rotation was plotted against the largest Republican dimension in Figure 2; this factor appears to be primarily an evaluative or "good-bad" dimension in both samples, and the correspondence between the two versions is quite marked. However, some subtle but quite reasonable distinctions did emerge. For example, Points 14, 19, and 16 (Roosevelt, Truman, and Stevenson) were apparently perceived as "more good" and Point 11 (McCarthy) as "more bad" by Democrats than by Republicans. On the other hand, Points 13, 10, 17, 3, and 2 (Nixon, MacArthur, Taft,

Dirksen, and Dewey) were evidently more highly evaluated by Republicans than by Democrats. The correspondence for the second, third, and fourth factors was even greater than for the first, the only marked difference appearing on a factor interpreted as a Republican vs. Democrat dimension, on which McCarthy was perceived as being "more typically Republican" by the Democrats.

The finding of only minor differences in the perceptual spaces of apparently diverse social groups is consistent with the results of previous studies in which multidimensional scaling structures were compared for groups having presumably different viewpoints about the stimuli. Abelson (1), for example, found only minor differences in the perception of attitude relationships between socialists and conservatives, as did Messick (3) between theological seminary students and military personnel. However, this consistent similarity of perceptual spaces across divergent groups is possibly a function of the scaling procedure and not necessarily a reflection of common perceptions. Since the scaling method for obtaining estimates of stimulus similarity involves the averaging of judgments over all the subjects in a particular group, consistent individual differences in the ratings, or different "points of view" about the stimuli, are blended together, and only the average conglomerate viewpoint is considered (4). Thus, the perceptual space of the "average Republican" might very well turn out to be extremely similar to that of the "average Democrat," but at the same time fail to reflect the range of consistent, and possibly quite different, viewpoints within either sample. A scaling technique is required which would first isolate empirically any consistent individual viewpoints about stimulus similarity and then provide for the derivation of separate multidimensional spaces for each viewpoint (cf. 8).

*Manuscript received: February 7, 1961*

*Revised Manuscript received: April 24, 1961*

Samuel Messick

Educational Testing Service

20 Nassau Street

Princeton, New Jersey

#### REFERENCES

1. Abelson, R. P., "A Technique and a Model for Multidimensional Attitude Scaling," *Public Opinion Quarterly*, 1954-1955, 18, 405-418.
2. Diederich, G. W., S. J. Messick and L. R. Tucker, "A General Least Squares Solution for Successive Intervals," *Psychometrika*, 1957, 22, 159-173.
3. Messick, S., "The Perception of Social Attitudes," *Journal of Abnormal and Social Psychology*, 1956, 52, 57-66.
4. Messick, S., "Dimensions of Social Desirability," *Journal of Consulting Psychology*, 1960, 24, 279-287.

5. Messick, S. and R. P. Abelson, "The Additive Constant Problem in Multidimensional Scaling," *Psychometrika*, 1956, 21, 1-15.
6. Torgerson, W. S., *Theory and Methods of Scaling*, New York: Wiley, 1958.
7. Tucker, L. R., "A Method for Synthesis of Factor Analysis Studies," PRS Report 984, Personnel Research Section PR 3075, Research Contract No. DA-49-083 OSA-116, Project No. 29530100 Subtask 75, Department of the Army, 1951.
8. Tucker, L. R. and S. Messick, "Individual Differences in Multidimensional Scaling," Princeton, New Jersey: Educational Testing Service Research Memorandum, 1960.

## Good Men, Bad Men, and the Perception of Reality

GWYNN NETTLER, *San Francisco, California*

With the rise of a scientific determinism and the associated decline of free will and sin as explanations of deviant behavior, it has become popular to look upon "badness" as sickness. Research designed to increase our knowledge of prostitution, radicalism, homosexuality, dictatorship, race prejudice, crime and delinquency often makes the assumption, implicitly and explicitly, that evil is illness.<sup>1</sup>

These assumptions have not gone unchallenged (10, 18, 35, 56, 65, 66), yet they flourish. Part of their viability flows from the "evil-causes-evil" fallacy (60, page 62), but these assumptions are sustained also in the vogue of psychoanalysis and by the ambiguities that reside in the idea of mental health, ambiguities into which prejudices have stepped. Attempts to define a "healthy personality" carry a load of personal preference and we are prone to call the man "sound" who behaves as we do, or as we like to think we behave (9). These preferences seem not merely idiosyncratic but also class-biased and ethnocentric.<sup>2</sup> They gather ideological support from the Platonic suggestion that goodness, truth, and beauty are the ultimate desirables and that they bear each other an intimate connection. Since this Platonism is part of the academic subculture in which most investigators have been steeped, there have been few dissident voices among sociologists (65) or humanists (37), although

<sup>1</sup> There is a tautological sense in which this is true: if illness is defined as anything bad that is caused.

For present purposes and euphony's sake I should like to use "evil," "bad," and "deviant" interchangeably, if it will be recognized that I do not believe that all deviation—even some of which offends society—is evil, and that what is evil according to classical standards includes much that is now acceptable.

This synonymy seems defensible in view of the orientation of *Social Problems* and *Social Disorganization* textbooks which, although they do not always say so directly, are "really" concerned with badness. And this remains true even when they are re-titled "Social Deviation", because it is only disapproved difference that one finds in the chapter headings. Thus one doesn't see nor expect discussions in these books of such deviations as being rich, powerful, beautiful, or creative, clear, and intelligent.

Thus saying, this paper will hereafter abandon the defense of placing quotation marks about these useful but fuzzy terms.

<sup>2</sup> For example, Witmer and Kotinsky (71) write that "... to be happy and responsible is to be healthy in personality." The words "happy" and "responsible" are so glorious that one is loath to question their relevance to mental health, but they are suspect simply because we do not expect a man to be happy regardless of circumstance and "responsibility" does not specify for whom nor how nor under what conditions with what limits.



psychologists seem increasingly reluctant to accept the vaunted relationship of health and goodness (19, 46, 59, 62, 63).

The hope that goodness is friendly to the truth, or, better, that knowing the truth will make us good, has been brewed into our definitions of the healthy man and the concoction then used to justify our values. The California study of the *Authoritarian Personality* is a prime example (1). Maslow makes the logic explicit:

The neurotic is not only emotionally sick—he is cognitively *wrong*! If health and neurosis are, respectively, correct and incorrect perceptions of reality, propositions of fact and propositions of value merge in this area, and in principle, value propositions should then be empirically demonstrable rather than merely matters of taste or exhortation (36).

If the scientist can show that bad behavior is a symptom of sickness and if a measure of this sickness is a perverted picture of the world, then the equation of morality (ours) with correctness seems established and certain happy consequences may be inferred, for example:

1. That truth and utility are *not* at war, as Pareto and the inconsistent Plato believed.
2. That science can "prove" values and thus give them the underpinning they lost with the decline of religious authority.
3. That, since evil-is-sickness-is-error, therapy remains what some like to think it is: truth-giving.
4. That we good people are also more factually correct.

But what if these assumptions are questionable? What if evil is a way of responding to an imperfect world by actors with limited resources in a determinate situation? What shall we say to such a sensitive writer as Han Su-yin (Dr. Elizabeth K. Comber) when she tells us, "So I'm not a good woman. If I'd been a good woman, where would I be now? I'll tell you—dead!"? Finally, what will it mean for our notions of psychotherapy and sociotherapy if, in those cases where evil people see a different world from that observed by good men, the evil eye should prove the clearer?

Raising such a possibility is itself an invidious task, but just such a one as the professional student of society must undertake as part of the work of describing how people behave. To the objector who would resist the question and assume its answer in advance of test, the reminder must be given that the conflict between truth and utility is an old one and that, while it remains unresolved, it is of crucial importance for any theory of personal or social improvement. It must also be considered that, while as scientists we protest our preference for truth over error, as culture-bound thinkers we may favor lies when their credentials are of the right sort.

Raising the uncomfortable question of the healthfulness of bad actors may place us, as meliorists, in a "bind" because there are realities which, correctly

perceived, are sickening, and we may have to choose between the "health" of the man who behaves badly because he sees accurately and the "health" of the man who behaves nicely because he has learned the popular ways of seeing falsely.

#### A PARTIAL EXAMINATION OF THE EVIL EYE

There is little doubt that bad people see the world differently from good. Many studies tell us so. Part of this perception involves preferences, tastes, self-conceptions and attitudes, but part of the difference in perception concerns how the world really is, and it is with this that we are concerned.

The differences in what good and bad people see "out there" share a generality because evil has been defined in socio-psychological research as deviation from a humanitarian, "progressive," middle-class, quasi-Christian,<sup>3</sup> Western-urban standard, so that it makes almost no difference which kind of badness one is inspecting—the kinds of belief about the world that allegedly differentiate the good and the bad man are of a persistent piece. In fact, one prominent study consciously explores the possibility that these evils are one morsel, as it attempts to relate political evil ("authoritarianism") to ethnic prejudice, crime, and psychic malfunction (1). What is crucial for present purposes is that these investigators believe the evil eye is astigmatic; they hold political evil (*their* conception) to be an ideology that fails as "... an objective appraisal of social reality, (and) tends to resemble a fantasy. . . ." (1, page 845).

The purportedly fantastic world seen by the bad man is a jungle. And a jungle may be defined as a place where strength and cunning decide and win—not "human-ness," "justice," "principle," nor any other ethic than force and fraud. It is a place where each against all is truer than all for each or, even, each for himself alone.

Since the perceptions reported as marking off the bad man from the good have this generality, it is not necessary to use the findings of every study, nor to limit our search to one kind of wrong, in order to examine the relative acuity of good and evil perception. A sample of such studies dealing with delinquency (6, 22, 23), juvenile narcotic users (14), fascism (1), and ethnic prejudice (1, 2) has been taken. From the measures which these studies report as differentiating, with some degree of probability, the attitudes of good and bad people, *all* statements relating to "external fact" as opposed to self-evaluation or personal behavior have been listed. The literature has then been scanned for evidence, tentative and partial though it might be, that would permit a

<sup>3</sup> The use of the qualifier, "quasi," is not flippant. Western society has moved a long way from the ideals and practices of a Christian community. For a statement of what the Christian mode would mean, see Jones (28) and Russell (54).

grading of these differentiating perceptions, using the bad person's view from his environment as the one to be tested, since it is he who is allegedly unhealthy and who eyes the world through need-distorted lenses.

In fairness to the investigators whose studies have been used, employment of their differentiating items should not be construed as meaning that these authors necessarily subscribe to the notion under question here—that badness is sickness, a symptom of which is a distorted perception of the world. The approach of certain of these investigators has been purely actuarial: what distinctions can be found that allow us to predict? But in other cases, notably those of Allport-Kramer and the Adorno group, the investigators do maintain that “good” perceptions are healthier and more accurate than evil ones.

The reality-perceptions of bad men, and one judge's estimate of their accuracy based on the sources cited, are listed below. To propose one's own estimate of perceptual accuracy is, of course, to run the risk of criticism *ad hominem* and the kind of “tele-psychoing” to which even behavioral scientists are not immune. However, this study has been entered with the explicit assumption that socio-psychological observers assessing good and evil conduct and mental health may themselves be wearing class-ethnic blinkers, so that it would hardly be pertinent to put this matter to their vote. Since the present concern lies more heavily with being heuristic than with being right, it would seem sufficient only to make a probable guess as to the state of the “real world” as good and bad people differentially see it, and to cite one's grounds for his estimate.

### *Some Available Evidence*

Ball (6) reports five statements that differentiate between delinquent and non-delinquent perceptions of the prevalence of stealing. The first reads:

*“How many people would steal if they had a good chance?”*

a. *The delinquent believes* “about half, most, or all.”

b. *Estimate:* Probably accurate; much depends on the meaning of “good chance.”

c. *Evidence:*

1. “. . . all employees of a chain-store were run through his (Keeler's) polygraph when the company complained it was losing more than one million dollars annually through petty thefts. Polygraph records indicated that fully three out of four employees were pilfering funds. This and subsequent experiences led Keeler to pronounce a rather cynical dictum generally held by lie-detection experts today: ‘65 per cent of people who handle money take money. . . .’” (17, pages 154-155)
2. “Lie detector tests of employees of certain Chicago banks showed that 20 per cent had taken money or property, and in almost all cases the tests were supported by confessions. Similarly, lie detector tests of a cross section sample of the employees of a chain store indicated that about 75 per cent had taken money or merchandise from the store.” (40)

3. "... the Comptroller of Currency reported that about three-fourths of the national banks examined in one period were violating the national banking laws and that dishonesty was found in 50.5 per cent of the national bank failures in the years 1865-1899, and 61.4 per cent in the years 1900-1919." (16, page 185)
4. "Undersecretary of the Treasury Fred A. Scribner, Jr., reported that a Treasury examination of 1956 income tax returns discovered that taxpayers had failed to report almost \$4,500,000,000. of interest and dividends received during that year. . . ." (55)
5. Investigators carefully dropped stamped, addressed postcards, letters, and letters bearing a lead coin simulating a 50-cent piece in various cities of the East and Midwest. Seventy-two per cent of the postcards were returned; 85 per cent of the blank letters; 54 per cent of the "coin-carrying" letters. "We conclude . . . that the public at large is very strikingly altruistic, manifesting obligingness, consideration, and responsibility. A sharp decline in the reliability of the public sets in under the effects of suggestion of financial gain. One-third of the altruistically minded are converted to selfish behavior. It is probable that an even larger proportion of the public at large is unreliable in such a financial matter." (41)
6. "Hume said one could be sufficiently sure about certain aspects of human nature to predict with accuracy what would happen to a quantity of gold left unguarded in a populous place. On the basis of human experience in a free society one can now say the same of anything, of however little value, that is portable. For in all our centers of dense population any possessions left unguarded on the doorstep—roller skates, bicycles, baby carriages, appliances, garbage cans, tools, lawn mowers, trash receptacles and the like—disappear as if by magic." (33)
7. *Reader's Digest* Survey, 1941:
  - Of 347 garages visited, 63 per cent were dishonest.
  - Of 304 radio shops, 64 per cent were dishonest.
  - Of 462 watchmakers, 49 per cent were dishonest (51).

A second of Ball's questions asks, "*Do you think many people have taken things at some time?*"

- a. *The delinquent* again believes that "about half, most, or all" have.
- b. *Estimate and Evidence*: There is little reliable evidence on this matter. However, one study of a non-criminal population is suggestive. Nettler (48) asked a heterogeneous sample of California residents these questions, among others, concerning their criminal conduct:
  - "Since age 16, how many times have you taken hotel towels or blankets as souvenirs?"
  - "Since age 16, how many times have you taken a newspaper from a stand without paying for it?"
  - "Since age 16, how many times have you taken something from a store without paying for it?"
  - "Since age 16, how many times have you kept money for yourself that belonged to someone else?"
  - "Since age 16, how many times have you taken another's property such as fruit, tools, library books, or other unsecured objects?"

A majority of the respondents admitted anonymously to one or more of these crimes.

Ball reports three other statements that distinguish good and bad perception: "Do you think many people would steal from their friends?" "Do you think many people would steal from a store if they had a good chance?" "... from a school?" Lack of evidence prevents a test of these items.

Chen (14) found that juvenile narcotics users tend to see a different social world than children who are free of this habit:

*"Most policemen treat people of all races the same."*

- a. *Narcotics user denies.*
- b. *Estimate:* probably an accurate perception—in the environment in which the question was asked.
- c. *Evidence:* (4; 7, pages 159, 164–165; 53; 60, pages 139–140; 64, page 95)

*"Most policemen can be paid off."*

- a. *Narcotics user agrees.*
- b. *Estimate and Evidence:* No one knows about "most policemen." But for evidence that this perception is not a fantasy, see (7, pages 245–246; 49, page 73; 60, pages 228–229, 383–389).

*"The police often pick on people for no good reason."*

- a. *Narcotics user agrees.*
- b. *Estimate and Evidence:* This must vary with race, class, and the police force. And no one knows, class by class, whether this is "often" true. But for evidence that this perception, again, is not a hallucination, see (7, page 750; 60, pages 331–341).

*"I am sure that most of my friends would stand by me no matter what kind of trouble I got into."*

- a. *Narcotics user denies.*
- b. *Estimate and Evidence:* Part of this is tautology; a "friend" is this kind of person.

Again, no one knows how much friendship there is in our society. And such as one finds probably varies socio-economically (13, pages 189–190; 34; 42, *passim*). "When you have no money, nobody takes an interest in you. I never had a friend. For six years it has been like that. We are living in filthy misery on the East Side—nobody can ever help us out of that." (25, page 17)

Jungles, by definition, are not friendly environments and we need to know their protean styles and their prevalence. It is more than one observer's judgment that friendship, as distinct from "cliqueship," is an abnormality on Manhattan's Seventh Avenue or San Francisco's Montgomery Street. And, if one talks as an industrial consultant to the \$52-a-week girls straitened desk-after-desk in the offices of Megalopolitan Insurance, he will not be reassured that many abide in the warmth of friendship. Or enter academia and observe the rarity of love and charity. Some participants have described the Academy as a genteel jungle, a microcosm of the competitive, political, envious and even unhappy world we usually associate with the strivers in gray flannel (8, 39, 69).

*"Everybody is just out for himself. Nobody really cares about anybody else."*

a. *Narcotics user* agrees.

b. *Estimate and Evidence*: As worded, this seems patently false. *Some* people care for others, at least in some milieux.

But this too must be class linked and the delinquent's perception may give an accurate picture of his world.

Such a statement poses, and leaves unanswered, three important questions about people's concern for others:

1. Under what circumstances does altruism develop?
2. How many people do care about how many others? What is the *range* of love by ethnic group and socio-economic status?
3. What is the *quality* of this concern? Love does not often come to us unalloyed. Portions of personal motivation, self-interest, ego are involved in altruism; the question is, "How much?" As Eric Hoffer has put it, "There is no doubt that in exchanging a self-centered for a selfless life we gain enormously in self-esteem."

As "concern for others" becomes a thing one does out of concern for himself—not to uphold self-principle, but to get something or somewhere—the delinquent's perception becomes more accurate. Perhaps he has put the emphasis on "*really* cares." Instance: involvement in a large community's welfare activities will impress one that many (how many?) of the civic-minded altruists do not "*really*" care for those they are helping. Their concern seems to this admittedly fallible observer sometimes to be a public expression that has status and recreational value, political, and even commercial value. Probably no one who has worked with charities and welfare agencies would dispute the existence (while he may debate the prevalence) of non-loving motives among board members, junior leaguers, service clubs, and other volunteer welfare-dispensers.

In the investigation by Hathaway and Monachesi (23) the following perceptions were found to distinguish predelinquent boys from others:

*"My mother was a good woman."*

a. *Predelinquent* denies.

b. *No evidence* available. A "good woman" is difficult to define.

*"One or more members of my family is very nervous."*

a. *Predelinquent* agrees.

b. *No evidence* in either direction.

*"The man who had most to do with me when I was a child (such as my father, stepfather, etc.) was very strict with me."*

a. *Predelinquent* agrees.

b. *Estimate*: probably true.



c. *Evidence:*

1. Glueck's "Social Factors" Prediction Table (21, p. 260):

	Weighted Failure Score:
#1. Discipline of Boy by Father—	
Overstrict or erratic	71.8
Lax	59.8
Firm but kindly	9.3

2. Anderson (3) reports that parents ranked by psychometrics as high in dominance and low in affection had aggressive and rebellious children.
3. "On every item in the interview in which intensity of punishment was rated, the fathers of the high aggression boys punished more severely than the fathers of the low aggression boys" (67). (This finding is qualified by the possibility that judges may have rated non-verbal punishment as more severe than verbal. Personal communication from the investigator).

*"When a man is with a woman he is usually thinking about things related to her sex."*

- a.
- Predelinquents*
- affirm.

- b.
- Estimate and Evidence:*
- This may be a function of age, but otherwise . . . ? It may also vary with status and culture. Much depends on the meaning of "with a woman" and "thinking about things." But, if the delinquent's perception is widely inaccurate, then Hollywood, the ladies' magazines, and American advertisers are also wearing distorting spectacles.

*"My parents have often objected to the kind of people I went around with."*

- a.
- Predelinquent*
- admits.

- b.
- Estimate:*
- probably true.

- c.
- Evidence:*
- Chwast (15) cites many studies that affirm the relationship between delinquent behavior and parental rejection/disapproval.

Gough and Peterson's items (22) distinguishing those predisposed to crime from those more immune are largely attitudinal and self-descriptive. However, there are two categories of statement that deal with how the world "out there" is perceived; three items are concerned with how "jungle" it is and nine with family life.

The predelinquent is depicted by Gough and Peterson as believing that: "I would have been more successful if people had given me a fair chance"; "Life usually hands me a pretty raw deal"; and "A person is better off if he doesn't trust anyone."

For all of these statements there is thus far no conclusive evidence. These items are, of course, both "objective" and "subjective"—they carry a heavy load of interpretation of what has happened to one—but, for present purposes, it need only be indicated that the predelinquent, particularly if he comes from the lower classes, is probably not distorting his reality greatly. And, in valida-

tion of the delinquent perception, one can cite all manner of learned non-delinquent. For example, Charles Stockard, the noted zoologist, believed, erroneously, that "... man is the only animal that gives deliberately false signals to his kind," and George Bernard Shaw has been quoted as saying:

"Man is the only animal of which I am thoroughly and cravenly afraid. I have never thought much of the courage of the lion-tamer. Inside the cage he is at least safe from other men. There is less harm in a well-fed lion. It has no ideals, no sect, no party, no nation, no class; in short, no reason for destroying anything it does not want to eat." (11)

Myers and Roberts (47) write of their class V (lowest status) patients, "They were reared in an environment where violence, aggression, hostility, and rebellion were accepted. (They) . . . received little warmth or affection."

David Riesman summarizes, "... the conditions of lower-class life do not encourage trust" (52).

The pre-criminal, according to Gough and Peterson, sees his family life as unsatisfactory and he is probably right. He believes:

*"My family has objected to the kind of work I do, or plan to do."*

*"My parents were too strict with me when I was a child."*

*"My parents never really understood me."*

*"My home life as a child was less peaceful and quiet than those of most other people."*

*"My parents have often disapproved of my friends."*

*"I was often punished unfairly as a child."*

And he denies that:

*"My home life was always very pleasant."*

*"My parents have generally let me make my own decisions."*

*"The members of my family were always very close to each other."*

The final form of the F (authoritarianism) scale devised by Adorno and associates (1) contains ten statements that describe how the world is or will be and to which the "fascist" subscribes. Of these ten items, one seems false, one true, and four might be correct depending upon the definitions adopted. For the remaining four statements no evidence is available.

*Probably false: "Wars and social troubles may someday be ended by an earthquake or flood that will destroy the whole world."*

*Probably true: "Human nature being what it is, there will always be war and conflict."*

See Hanson Baldwin (5) or Freud (20).

*Possibly true, depending upon the definitions given:*

*"People can be divided into two distinct classes, the weak and the strong."*

As with all defining statements, the "validity" of this one is a function of someone's purposes in making such a classification. While field theorists warn against such Aristotelian dichotomies, many scientists are convinced of the usefulness of dichotomous assumptions for certain purposes. See, for example, George Kelly (30).

*"A person who has bad manners, habits, and breeding can hardly expect to get along with decent people."*

As conventionally understood, this is probably a true statement; much depends on who one thinks is "decent."

*"Most of our social problems would be solved if we could somehow get rid of the immoral, crooked, and feeble-minded people."*

Since there is no agreed-upon roster of social problems, it is impossible to say how many problems would constitute "most." However, see any social problems text.

*"The wild sex life of the old Greeks and Romans was tame compared to some of the goings-on in this country, even in places where people might least expect it."*

Comparative study of this nature seems insurmountably difficult, although challenging. See almost any issue of *Playboy*, *Confidential*, or the daily press. Also, Ben Hecht (24).

*No evidence available:*

*"If people would talk less and work more, everybody would be better off."*

*"Most people don't realize how much our lives are controlled by plots hatched in secret places."*

The words "plot" and "secret" make the statement difficult to evaluate; otherwise see Hunter (26) and Mills (42).

*"The businessman and the manufacturer are much more important to society than the artist and the professor."*

*"Familiarity breeds contempt."*

Most of the items in the E (Ethnocentrism) scale devised by the same group of investigators (1, page 142) are too attitudinal and ambiguous to permit test, but one seems quite probably false, and two possibly true:

*Probably false: "The people who raise all the talk about putting Negroes on the same level as whites are mostly radical agitators trying to stir up conflicts."*

*Possibly true: "The best guarantee of our national security is for America to have the biggest army and navy in the world and the secret of the atom bomb."*

This item is dated. If its intent is read as of 1961, our government seems to be acting as if preparedness were "the best guarantee," even though American sociologists may not feel too secure in it.

Again, see Baldwin: "The right fist of all-out nuclear power must remain ready as a sanction to help ensure limited war" (5, page 43).

Or, Morgenstern: "The problem, the great problem, of our military planners is to organize and maintain armed forces capable of fighting any kind of war anywhere" (43).

*"The trouble with letting Jews into a nice neighborhood is that they gradually give it a typical Jewish atmosphere."*

This offensive statement is emotionally loaded by the phrase, "The trouble with . . ." and the adjective, "nice." If one discounts these attitudinal indicators and

restricts the sentence to its factual representation, to the possibility of ethnic invasion and succession, it seems true in the larger metropolitan areas of the United States (71).

*Ambiguous and No Evidence Available:*

*"One trouble with Jewish businessmen is that they stick together and prevent other people from having a fair chance in competition."*

*"There may be a few exceptions, but in general Jews are pretty much alike."*

*"There is something different and strange about Jews; it's hard to tell what they are thinking and planning, and what makes them tick."*

*"Manual labor and unskilled jobs seem to fit the Negro mentality and ability better than more skilled or responsible work."*

*"Most Negroes would become overbearing and disagreeable if not kept in their place."*

*"The worst danger to real Americanism during the last 50 years has come from foreign ideas and agitators."*

*"America may not be perfect, but the American way has brought us about as close as human beings can get to a perfect society."*

Allport and Kramer (2), testing Ichheiser's psychoanalytic hypothesis about anti-Semitism (27), find that anti-Semites are differentiated from the unprejudiced by their greater fear of swindlers than of gangsters. They say, "... those who are more afraid of swindlers also have higher prejudice scores in general. ( $\chi^2 = 7.17$ , significant at the 1 per cent level). Those who are more afraid of gangsters (may we not say, a more natural and normal fear?) are less prejudiced" (2, pages 33-34).

Allport and Kramer may choose to call the good man's greater fear of gangsters "natural and normal," whatever that means, but such honorific labelling does nothing to root this fear in a clearer perception of reality. Again, the good man, Allport and Kramer included, seems to be looking at the world through distorting spectacles, for the evidence is abundant that good and bad people alike are at least as apt to be defrauded by swindlers as bludgeoned by gangsters. Barnes and Teeters (7, page 69) call it a toss-up, but Nietzsche (32), White (68), and Bernard (12, page 36), among others, see fraud as the more common danger. Sutherland felt it probable that "... fraud is the most prevalent crime in America" (60, page 42), Schur calls swindling, "... a strongly entrenched national phenomenon. . . ." (57, page 269), and Nettler (48) found "crimes of deceit" to be admitted anonymously by a majority of a non-criminal population. Reference to works on white-collar crime (61) and confidence games (38) confirms the suspicion that Allport-Kramer and their unprejudiced subjects are less able to test the reality about them (in this department at least) than their evil respondents.

*A Tally*

It appears that good people see the world at least as inaccurately—to put it mildly—as bad ones. Any "score" here is subject to criticism and can only

be offered tentatively. For, aside from the quarrels with one investigator's assessment of reality, even where he cites "evidence" for his view, there remains the possibility that good and bad men reside in different "real" worlds and, hence, where they differ they may be equally correct or incorrect. For this reason Table 1 presents three tallies: one of all 48 items as judged for approximation to actuality in the text above, and two tallies—a "rough" and a "soft" one—of the 37 items that remain after the family history items in the Gough-Peterson and Hathaway-Monachesi studies are discounted.

A reading of these 37 non-personal items shows that the possible residence of good and bad men in different "real" worlds does not explain all the differ-

TABLE 1

*A Tally of the Relative Accuracy of the Perceptions Of Good and Bad Men*

## TALLY I: ALL ITEMS

	Study							Total
	Ball	Chein	H-M	G-P	F-Scale	E-Scale	A-K	
Bad Man Accurate	2	5	3	9	5	2	1	27
Good Man Accurate	0	0	0	0	1	1	0	2
Evidence Lacking or Ambiguous	3	0	2	3	4	7	0	19
Total	5	5	5	12	10	10	1	48

TALLY II: FAMILY HISTORY ITEMS OMITTED  
"ROUGH" VERSION

Bad Man Accurate	2	5	1	0	5	2	1	16
Good Man Accurate	0	0	0	0	1	1	0	2
Evidence Lacking or Ambiguous	3	0	2	3	4	7	0	19
Total	5	5	3	3	10	10	1	37

TALLY III: FAMILY HISTORY ITEMS OMITTED  
"SOFT" VERSION

Bad Man Accurate	1 <sup>a</sup>	1 <sup>b</sup>	1	0	1 <sup>c</sup>	1 <sup>d</sup>	1	6
Good Man Accurate	0	0	0	0	1	1	0	2
Evidence Lacking or Ambiguous	4	4	2	3	8	8	0	29
Total	5	5	3	3	10	10	1	37

<sup>a</sup> Calls "many people have taken things" lacking in evidence.

<sup>b</sup> Considers only the delinquent negation of "Most policemen treat people of all races the same" as an accurate perception in the environment in which the question was asked.

<sup>c</sup> Calls the four items judged "possibly true" lacking in evidence.

<sup>d</sup> Calls the item about "Jews in nice neighborhoods" ambiguous and lacking in evidence.

ences in perception; a generalized view of the shared world remains and, from the items and evidence at hand, it cannot be concluded that the evil eye sees it poorly.

#### DISCUSSION

It has been argued that the conception of evil action as sickness is questionable. Particularly doubtful is the assumption that a prime symptom of the illness that generates badness is a false picture of the world. If this now seems obvious, one wonders how such an idea could have been given professional credence. A few sources of this intellectual infection may be suggested.

It would seem that the notion that bad people must be sick is the spawn of American pragmatic optimism (All problems are soluble if we but think; do something!) bred with the classic values of goodness, truth, and beauty in an atmosphere in which the prestige of religious explanation has been declining while that of popular psychologizing has been rising.

Having "advanced" beyond blaming the bad man for his moral depravity, the middle-class investigator proposes to treat him for his sickness. This proposal is emboldened by the optimistic assumption that goodness and health (which includes telling, seeking, and seeing the truth) are reciprocally related. With faith so set, it follows that evil may be cured, like other infirmities, and that an important part of the cure lies in the bad, sick man's taking psychotherapeutic exercises in correct perception—of what he has done, and why, and of how people "really" are as opposed to what he thought they were.

This tentative study calls into question this popular thoughtway. Bad actors may not be sick—at least no more so than good men—and, particularly, they may not suffer from perverted perception. Rather, if the present tally has validity, it is the good man who sees relatively inaccurately.

Two questions persist:

—How jungle-like is our society?

—To what extent is behavior a function of how one perceives his social world?

On the first question our "authorities" give conflicting answers. Freud once advised that "Men are not only worse, but also better, than they think they are," but, another time, he writes, "*Homo homini lupus*; who has the courage to dispute it in the face of all the evidence in his own life and in history?" (20)

A philosopher reviewing a work on "New Knowledge in Human Values" asks its psychologist-editor, "If Maslow is right in saying that 'our deepest needs are *not* . . . dangerous or evil or bad,' why are our surface appetites often so?" (58)

And Ortega tells us, "The very name, 'society,' as denoting groups of men who live together, is equivocal and utopian." (50)



Relative to the second question, disputes persist concerning the value of accurate perception (44).

- "It is not wholesome to live by illusion," Gordon Allport assures us.
- "Mankind cannot stand too much reality," T. S. Eliot replies.
- "The golden virtues are "... love, truth, beauty, self-realization . . . (they) conduce to psychological health," Huston Smith retorts.
- "But "... mental integrity may rest on the capacity for denial, for sustained repression of truth," answers Philip Roche.

If more thorough investigations support present assumptions about the likely answers to our two questions, the result will be unfortunate for those of us who have learned to value truth. With the Existentialists, we may have to agree that some truths are sickening. And with Plato and the many religionists, we may have to agree that men need the control as well as the comfort of myth.

*Manuscript received: January 30, 1961*

*Revised Manuscript received: April 26, 1961*

Gwynn Nettler

210 Post Street, #915

San Francisco 8, California

#### REFERENCES

1. Adorno, T. W., et al., *The Authoritarian Personality*, New York: Harper, 1950.
2. Allport, G., and B. M. Kramer, "Some Roots of Prejudice," *Journal of Psychology*, 1946, 22, 9-39.
3. Anderson, J. P., "A Study of the Relationships Between Certain Aspects of Parental Behavior and the Attitudes and Behavior of Junior High School Students," *Contributions to Education*, #809, New York: Teachers College, Columbia University, 1940.
4. Axelrad, Sidney, "Negro and White Institutionalized Delinquents," *American Journal of Sociology*, 1952, 57, 569-574.
5. Baldwin, H. W., "Limited War," *The Atlantic Monthly*, May, 1959, 203, 35-43.
6. Ball, J. C., "Delinquent and Non-Delinquent Attitudes Toward the Prevalence of Stealing," *Journal of Criminal Law, Criminology, and Police Science*, 1957, 48, 259-274.
7. Barnes, H. E., and N. K. Teeters, *New Horizons in Criminology*, New York: Prentice-Hall, 1951.
8. Barr, Stringfellow, *Purely Academic*, New York: Simon and Schuster, 1958.
9. Barron, Frank, "Towards a Positive Definition of Psychological Health," paper read to the American Psychological Association, 1955.
10. Barron, M. L., *The Juvenile in Delinquent Society*, New York: Knopf, 1954.
11. Basso, H., "Foxy Like a Grandpa," *New Yorker*, 1949, 25, 109.
12. Bernard, L. L., *Social Control*, New York: Macmillan, 1939.
13. Broom, L., and P. Selznick, *Sociology*, Evanston: Row-Peterson, 1955.
14. Chein, I., "Narcotics Use Among Juveniles," *Social Work*, 1956, 1, 50-60.

15. Chwast, J., "Perceived Parental Attitudes and Pre-Delinquency," *Journal of Criminal Law, Criminology, and Police Science*, 1958, 49, 116-128.
16. Cressey, D. R., *Other People's Money*, Glencoe: Free Press, 1953.
17. Deutsch, A., *The Trouble with Cops*, Boston: Little, Brown, 1950.
18. Frank, L. K., *Society as the Patient*, New Brunswick: Rutgers University Press, 1948.
19. Freides, D., "Toward the Elimination of the Concept of Normality," *Journal of Consulting Psychology*, 1960, 24, 128-133.
20. Freud, S., *Civilization and Its Discontents*, New York: Doubleday, 1953, p. 61.
21. Glueck, S., and E. Glueck, *Unraveling Juvenile Delinquency*, Cambridge: Harvard University Press, 1950, pp. 257-262.
22. Gough, H., and D. R. Peterson, "The Identification and Measurement of Pre-dispositional Factors in Crime and Delinquency: A First Report," paper read at the 1951 meeting of The Western Psychological Association.
23. Hathaway, S. R., and E. D. Monachesi, "The Personalities of Predelinquent Boys," *Journal of Criminal Law, Criminology, and Police Science*, 1957, 48, 149-163.
24. Hecht, B., "Sex in Hollywood," *Esquire*, 1954, 41, 35 and *passim*.
25. Hellersberg, E. F., *The Individual's Relation to Reality in Our Culture*, Springfield: Thomas, 1950.
26. Hunter, F., *Top Leadership, U.S.A.*, Chapel Hill: University of North Carolina Press, 1959.
27. Ichheiser, G., "Fear of Violence and Fear of Fraud, with Some Remarks on the Social Psychology of Anti-Semitism," *Sociometry*, 1944, 7, 376-383.
28. Jones, W. T., *A History of Western Philosophy*, New York: Harcourt, Brace, 1952, pp. 342-343.
29. Kahn, Jr., E. J., "Annals of Crime," *New Yorker*, 1959, 35, 122-153.
30. Kelly, G. A., *The Psychology of Personal Constructs*, New York: Norton, 1955, vol. I, pp. 109-110.
31. Lambert, W. W., L. M. Triandis, and M. Wolf, "Some Correlates of Belief in the Malevolence and Benevolence of Supernatural Beings: A Cross-Societal Study," *Journal of Abnormal and Social Psychology*, 1959, 58, 162-169.
32. Levy, O. (ed.), *The Complete Works of Friederich Nietzsche*, vol. 7, *Human, All-too-Human*, London: Foullis, 1911, pp. 200-201.
33. Lundberg, F., *The Treason of the People*, New York: Harper 1954, p. 289.
34. Lynd, R. J., and H. Lynd, *Middletown*, New York: Harcourt, Brace, 1929, p. 272.
35. Masling, S. M., "How Neurotic is the Authoritarian?" *Journal of Abnormal and Social Psychology*, 1954, 49, 316-318.
36. Maslow, A. H., *Motivation and Personality*, New York: Harper, 1954, p. 204.
37. Maugham, S., *The Summing Up*, New York: Doubleday, 1938.
38. Maurer, D. W., *The Big Con*, New York: Bobbs-Merrill, 1940.
39. McCarthy, M., *The Groves of Academe*, New York: Harcourt, Brace, 1952.
40. McEvoy, F. P., "The Lie-Detector Goes Into Business," *Reader's Digest*, 1941, 38, 69-72.
41. Merritt, C. B. and E. G. Fowler, "The Pecuniary Honesty of the Public at Large," *Journal of Abnormal and Social Psychology*, 1948, 43, 90-93.
42. Mills, C. W., *The Power Elite*, New York: Oxford University Press, 1956.
43. Morgenstern, O., *The Question of National Defense*, New York: Random House, 1959.
44. Moore, W. E., and M. M. Tumin, "Some Social Functions of Ignorance," *American Sociological Review*, 1949, 14, 787-795.
45. Morris, C. W., *The Variety of Values*, Chicago: University of Chicago Press, 1956.

46. Mowrer, O. H., "Sin, the Lesser of Two Evils," *American Psychologist*, 1960, 15, 301-304.
47. Myers, J. K., and B. H. Roberts, *Family and Class Dynamics in Mental Illness*, New York: Wiley, 1959, page 253.
48. Nettler, G., "Antisocial Sentiment and Criminality," *American Sociological Review*, 1959, 24, 202-218.
49. Nossiter, B. D., "The Teamsters: Corrupt Policemen of an Unruly Industry," *Harper's*, 1959, 218, 70-76.
50. Ortega y Gasset, Jose, *Concord and Liberty*, New York: Norton, 1946, p. 24.
51. Reader's Digest, 1941, 38, July, August, September issues.
52. Riesman, D., "Political Communication and Social Structure in the United States," *Public Opinion Quarterly*, 1956, 20, 60.
53. Robison, S. M., *Juvenile Delinquency: Its Nature and Control*, New York: Holt-Dryden, 1960, p. 211.
54. Russell, B., *Why I Am Not a Christian*, New York: Simon and Schuster, 1957.
55. Sahlman, H., "Taxes: Evasion and Avoidance," *Commentary*, 1959, 28, 447-448.
56. Schuessler, K. F., and D. R. Cressey, "Personality Characteristics of Criminals," *American Journal of Sociology*, 1950, 55, 476-484.
57. Schur, E. M., "Sociological Analysis of Confidence Swindling," *Journal of Criminal Law, Criminology, and Police Science*, 1957, 48, 296-304.
58. Smith, H., review of A. H. Maslow (ed.), "New Knowledge of Human Values," *Saturday Review*, 1959, 42, 25.
59. Smith, M. B., "Research Strategies Toward a Conception of Positive Mental Health," *American Psychologist*, 1959, 14, 673-681.
60. Sutherland, E. H., and D. R. Cressey, *Principles of Criminology*, Philadelphia: Lippincott, 1955.
61. Sutherland, E. H., *White Collar Crime*, New York: Dryden, 1949.
62. Szaz, T. S., "The Myth of Mental Illness," *American Psychologist*, 1960, 15, 113-118.
63. Szaz, T. S., "Naming and the Myth of Mental Illness," *American Psychologist*, 1961, 16, 59-65.
64. Taft, D. R., *Criminology*, New York: Macmillan, 1950.
65. van den Haag, E., "Psychoanalysis and Its Discontents," in S. Hook (ed.), *Psychoanalysis, Scientific Method, and Philosophy*, New York: New York University Press, 1959.
66. Volkman, A. P., "Delinquent and Nondelinquent Personality," *Social Problems*, 1958-59, 6, 238-245.
67. Walder, L. O., "An Attempt at an Empirical Test of a Theory," in *The Application of Role and Learning Theories to the Study of the Development of Aggression in Children*, Washington: American Psychological Association, 1958.
68. White, L. A., *The Evolution of Culture*, New York: McGraw-Hill, 1959, pp. 346-347.
69. Williams, G., *Some of My Best Friends Are Professors*, New York: Abelard-Schuman, 1958.
70. Wirth, L., *The Ghetto*, Chicago: The University of Chicago Press, 1929.
71. Witmer, H. L., and R. Kotinsky, *Personality in the Making*, New York: Harper, 1952, p. xviii.

## Group Morale and Competence of the Leader <sup>1</sup>

ROBERT L. HAMBLIN, KEITH MILLER, AND JAMES A. WIGGINS,  
*Washington University*

Inasmuch as low morale may be reflected in low productivity, high turnover, strikes, and other organizational as well as individual costs, it is not surprising that the maintenance of morale is a continuous concern in centralized organizations, particularly bureaucracies. Nor is it surprising that morale has been a continued focus of organizational theory and research.

Since World War II the theory and research on morale in organizations have been influenced considerably by the concepts and hypotheses used by Kurt Lewin (11, 12). These hypotheses have been investigated over and over again both in the laboratory and in the field, with cumulative results that generally support the conclusion that democratic leadership which involves decision-making by organization members increases morale, whereas autocratic leadership which involves centralized decision-making decreases morale. In addition, a number of these studies suggest that centralized decision-making lowers morale by removing the opportunity for initiative from organization members (2, 10). In results, the prolific efforts of Lewin, his students, and his admirers have contributed much to an understanding of the sources of morale and its loss in organizations.

However, the influence of Lewin and his colleagues has been disfunctional in one way: the success of their hypotheses in research has drawn attention away from other equally classical hypotheses <sup>2</sup> which are relevant to variations in morale in organizations. In particular, the morale-competence hypothesis, which was proposed independently by both Cooley (3) and Simmel (17), has been neglected.

The hypothesis is simply that morale in an organization is a function, partially at least, of the relative technical competence of the leader in coping with the organization's problems. In other words, if the leader, the member of the organization with decision-making authority, is less competent than other members, then morale will be low, whereas if he is more competent, morale will be high. The morale-competence hypothesis has been neglected at least

<sup>1</sup> This investigation was supported by a grant from the Office of Naval Research, Contract Nonr 816 (11). One version of this paper was issued as Technical Report #4 to the Office of Naval Research, and a second version was read at the 1960 meetings of the American Sociological Association. The authors are associated with the Social Science Institute and the Department of Sociology-Anthropology.

<sup>2</sup> For an early discussion of democratic-autocratic hypotheses, see Cooley (3, pp. 262-263).

to the extent that it has never been tested intentionally. However, results which are consistent with the hypothesis have emerged as incidental findings in three separate investigations.

First, Katz and Kahn (9) report a factor analysis of the various dimensions of morale of workers in industry. Two of the questions most highly loaded on the factor measuring satisfaction with the foreman involve an evaluation of the overall competence of the foreman, .75, and his relative knowledge of jobs in the work group, .65. In another factor, involving satisfaction with the company, a question assessing the overall competence of management was more highly loaded than any other question, .75. Thus, these factors give some indication that morale is associated with the competence of leaders in industrial organizations. Unfortunately, morale was measured inferentially; no questions were included which directly measured satisfaction with foreman and satisfaction with management.

Secondly, in a study of leadership, motivation, and attitudes of research teams, Baumgartel (1) found that subordinate scientists' morale, in terms of a sense of progress toward the scientific goals, was strongly associated with the scientific competence of the research leader. Finally, Maier (13), in a laboratory study comparing the productivity of authoritarian groups with skilled and with unskilled leaders, found incidentally that the skilled leaders are more frequently accepted by the group.

Since these findings were more or less fortuitous, it is inappropriate to think of any one of the studies as a rigorous test of the morale-competence hypothesis. However, taken together, the findings from the three investigations lend considerable support to the hypothesis—at least enough to warrant the more interesting investigation of questions of theory. Why should a relationship obtain between morale and competence? Three rather different lines of reasoning may be used to derive the relationship. The first centers around the concept of reinforcement, the second around the concept of reciprocity, and the third around the concept of conflict.

Although expressed differently in the two fields, the first theory has roots deep in both psychology and sociology. In the context of psychological theory, morale may be conceptualized as a manifestation of collective, generalized, positive reinforcement. Generalized reinforcement occurs when something—say, a recurring event—is usually linked with reinforcement. When generalized, the event elicits the response usually elicited by reinforcement. If members are usually rewarded or reinforced positively for participating in the organization, their feelings toward their situation in the organization will be the result of the generalization of this reinforcement. Therefore, morale, as a collective result of generalized reinforcement, will vary directly with reinforcement.

In sociology, one version of this last theorem was used in explaining variations in the morale of American soldiers. Merton (15), among others, discovered that in order to understand the relationship of morale to *negative* reinforcement—or, in his terminology, to deprivation—it is necessary to measure relative rather than objective deprivation. Based on a large number of findings, the conclusion was that morale varies inversely with relative deprivation.<sup>3</sup>

Through the use of this conclusion, the morale-competence hypothesis may be derived as follows: Assume that morale varies inversely with relative deprivation or directly with relative positive reinforcement. Assume also that reinforcement varies directly with organizational success or productivity. Assume further that organizational success or productivity is related to the competence of the leader. Then, to the extent the leader is the most competent group member, productivity and success will be maximized as will reinforcement, and consequently, morale will be high. On the other hand, to the extent the leader is incompetent, group productivity will suffer, reinforcement will be low, and morale will be low and resentment high.

Note that the focus in this derivation is not competence but the fact that the competent leader helps the organization to be more successful, and that the success of the organization results in reinforcement for organizational members. Thus, reinforcement due to the leader is the "causal" variable for predicting morale when this theory is used.

The second theory, as expressed by Cooley (3, pp. 180–181), is as follows:

It is felt . . . that the differences among men should be functional and intrinsic, not arbitrary or accidental. The sense of justice is usually strong among the members of a . . . group, the basis for determining what is just being the perception of some purpose which every one is to serve, each in his own way, so that he who rightly holds a higher place is the one who can function best for the common good. It does not hurt my self-respect or my allegiance to remain a common seaman while another becomes captain of the ship, provided I recognize that he is the fitter man for the place; and if the distribution of stations in society were evidently of this sort there would be no serious protest against it.

As mentioned above, Lewin's colleagues found that the curtailment of initiative lowers morale. Cooley suggests here that this relationship is conditional—that it does not occur when the leader is perceived as being relatively more competent than other organization members. His prediction is based on a "sense of justice" or, more sociologically, on the norm of reciprocity<sup>4</sup> which is violated unless "he who rightly holds a higher place is the one who can function best for the common good"—that is, unless he who is assigned the prerogatives and advantages of authority is the one who is perceived as being

<sup>3</sup> For an early investigation which established the importance of *relative* reinforcement in psychology, see Tinklepaugh (19).

<sup>4</sup> For a discussion of the norm of reciprocity, see Gouldner (7).



able to give the most in guiding organizational members in the achievement of the organizational goals. To the extent that a reciprocity between assigned authority and perceived competence obtains during a given period, morale will be high; to the extent it does not, morale will be low.<sup>5</sup>

If we combine Cooley's and Lewin's theoretical frames of reference, a third derivation may be made using conflict as the mediating variable rather than reciprocity. In an organization where a leader makes the decisions, morale may be negatively related to the number of suggestions made by organizational members which conflict with the leader's suggestions and eventual decisions. The number of conflicting suggestions by organization members may, in turn, be negatively related to the perceived competence of the leader. When the leader is perceived as being incompetent, organization members may well resent giving up their initiative to the leader because they may feel that they themselves can do a better job. Because they feel it is necessary for organizational success, members may offer many suggestions which conflict with the leader's. However, when the leader is perceived as being more competent than the others, organizational members may readily give up initiative because they know the leader can wield it better than they themselves can. In this case, because they do not feel it is necessary for organizational success, organization members may not offer suggestions which conflict with the leader's. Thus, the third derivation may be expressed as follows: if morale is negatively related to the number of conflicting suggestions and if the number of conflicting suggestions is negatively related to perceived competence, then morale will be positively related to perceived competence.

It may be noted that in the theories, three different developmental sequences are postulated, the first involving at least four hypotheses, the second perhaps two, and the third involving two hypotheses in addition to those postulated in the second theory. One way of testing the validity of these theories is actually to test the several hypotheses postulated in each of the developmental sequences. Although ultimately such a complete test is desirable, a test of only three hypotheses is critical to the evaluation of the three theories. The three hypotheses—one from each theory—in each case involve morale and a "causal" variable, the variations of which are postulated to produce variations in morale. In the first theory, the critical hypothesis involves morale and relative reinforcement, in the second, morale and perceived competence of the leader, and in the third, morale and conflicting suggestions by members.

<sup>5</sup> As this paper was going to press, it came to our attention that Durkheim, some nine years before Cooley, noted the relationship between morale and competence—or, using his terms, solidarity and talent—in the next-to-last chapter of his doctoral dissertation, *The Division of Labor* (6, pp. 374–381). Like Cooley, he explained the relationship in terms of justice.

The tests of these hypotheses are critical because these hypotheses allow us to ascertain the amount of the variance of morale actually controlled by the "causal" variables, and with it the amount of the variance of morale associated with each of the three developmental sequences. In other words, the first theory is important to the extent relative reinforcement controls the variance of morale; the second to the extent perceived competence controls the variance of morale; and the third to the extent the number of conflicting suggestions controls the variance of morale. It is quite possible that all three of the theories may be important, or any two, or that only one or perhaps none may be important. However, since the critical hypothesis in the first theory is well supported by research in social science, the expectation is that the results will show that this theory is important. The question is, will the derivations based on the critical hypotheses in the second and third theories also be supported? Will perceived competence of the leader, on the one hand, and conflicting suggestions by organization members, on the other, exert any important influence on morale independent of each other and independent of the influence of relative reinforcement? The remainder of this report consists of the methods and results of an experiment designed to test the three critical hypotheses and thus to evaluate the three theories which relate morale and the competence of the leader.

#### METHOD

Inasmuch as authority and leadership involve decision-making, an experimental situation was developed which in essential ways would simulate decision-making in organizations. Although most organizations are large, the aspects of decision-making which are apparently relevant to the three theories outlined above are present in small groups.<sup>6</sup> Therefore, for convenience and in order to have more than one non-leader, three-person groups were used.

The sample consists of 42 male college students who, together with one confederate, composed 21 three-person groups. From all curricula, the subjects were obtained through the cooperation of the Army and Air Force R.O.T.C. Members of each group were visually and physically separated, as each was ushered into one of three experimental rooms upon arrival. The only communication was carried on by means of an electronic intercom system which allowed full verbal communication between the three experimental rooms and a control room used by the experimenter.

In order to test the theories, it was necessary to operationalize the variables which are essential in predicting morale on the basis of three theories: rein-

<sup>6</sup> Note that, in the investigations cited above, the hypotheses worked equally as well in industrial work groups, large-scale industrial organizations, scientific research teams, and small laboratory groups.

FIGURE 1  
*The Choice Matrix*

	A	B	C	D	E	F	G	H
Yellow	9	8	16	10	-5	-6	-7	-4
Green	20	8	-6	-20	8	5	-6	-9
Brown	11	-3	13	-7	18	-5	11	-4
Red	12	-10	-8	13	15	-7	-7	7
Blue	-7	-9	12	10	5	13	-12	-6
Orange	-11	13	-10	8	9	-8	8	-7
Purple	-6	11	13	-13	13	-7	-5	20

forcement because of the success of the leader, perceived competence of the leader, and conflicting suggestions. However, for the tests to be possible, reinforcement, perceived competence, and conflicting suggestions could be but moderately associated, at most. Consequently, a situation was designed in which, much of the time, success and the resulting reinforcement would be perceived by group members as being attributable to the leader's luck rather than to his competence, and in which conflicting suggestions were allowed to occur as they might. The task finally selected was a gambling task involving the choice matrix in Figure 1.<sup>7</sup> For ease of communication, each of seven rows was identified by a color and each of eight columns by a letter from A to H inclusive. The cells of the matrix contained numbers which varied between plus and minus 20.

The subjects were told that over a series of ten trials the experimenter would choose various columns of the matrix according to a system which

<sup>7</sup> We wish to thank Edwin P. Hollander, whose similar use of a choice matrix gave us the idea that crystallized the elements of this design.

would not be disclosed. Before each column selection was announced by the experimenter, the subjects were allowed two minutes for discussion to select a row. They were then awarded or fined a sum of money corresponding to the plus or minus entry in the cell at the intersection of the column chosen by the experimenter and the row chosen by the subjects. Thus, their real tasks were (a) to determine the experimenter's system for choosing columns or letters, and (b) to choose that row which would maximize their winnings in the column which the experimenter chose. The subjects could maximize their winnings only if they could discover the experimenter's system and thereby predict his selection.

Before the task began, each group was told that to guarantee a decision a leader would be appointed by a random method: the rooms would be assigned according to the order of appearance at the experiment and the leader would be designated by a sign on his desk. The leader was then asked to identify himself. In every case, of course, the leader turned out to be the confederate, and he was given full and final authority to make each of the ten decisions, while the subjects were asked simply to take adviser roles.

Each group began the task with \$1.00 credit; the members were told that each would be awarded one-third of the cumulative earnings. After each of the ten row-selections by the group, the experimenter announced his column choice, the amount they won or lost, and the amount of their cumulative earnings. The manipulation was achieved by the simple expedient of having the experimenter actually delay his column selection until the group had made their decision. With the values in the cells arranged as they were, given any initial row-selection by the subjects and a subsequent disagreeing selection by the leader, the experimenter could pick a column which would show a positive sum of money for one and a negative sum for the other. In this way the experimenter could control the number of times the subjects would have made a winning decision, and the number of times the leader actually did make it. The number of times the leader made a decision against the advice of his subjects was controlled at six in all three experimental variations which were used.

The experimental manipulation included three experimental conditions involving high, medium, and low success of the leader. In the high success variation the leader made a winning decision eight times and a losing decision twice, while the followers advised winning decisions twice and losing decisions eight times. In the medium condition, the leader made five winning and five losing decisions and the followers advised five losing and five winning decisions. In the low success condition, the leader made eight losing and two winning decisions, while the followers advised eight winning and two losing decisions.

With the use of these procedures, success did not necessarily reflect the leader's competence. Frequently and quite randomly, successful decisions occurred because of what must have appeared to the subjects to be pure chance. Often the column selected by the experimenter was different from the column predicted by the leader and yet the value of the critical cell at the intersection of the column selected by the experimenter and the row selected by the leader would still be positive, indicating success, since a sum would be added to the winnings.

Consequently, appearing somewhat competent and not just lucky was a major problem for the leader. A number of techniques were developed for the leader to maintain a semblance of competence. The leader often chose two alternatives; this markedly increased the probability of being right as well as successful. Other times the leader would admit failure and place the blame on his improper use of his "system" rather than on the "system" itself. Having a "system" that worked seemed to be necessary if success was to be perceived as being more than luck. However, this tendency to admit and rationalize failure was uncontrolled.

Thus, relative success was manipulated experimentally somewhat independently of what we hoped would appear to be competence. Then the critical independent variables—reinforcement, perceived competence, and conflict—were measured.

First, relative reinforcement was measured simply by subtracting from the amount the group actually won what the group would have won, had the subjects' original suggestions been adopted by the leader. Secondly, the competence of the leader as perceived by the subjects was measured in a post-experimental questionnaire by having the subjects evaluate how intelligently the leader used the ideas of the group during the experiment:

How do you feel about the leader's use of the ideas you and the other participant gave? Circle the best alternative; "X" the next best.

- a) Very intelligent use
- b) Quite intelligent use
- c) Somewhat intelligent use
- d) Not at all intelligent use
- e) Somewhat unintelligent use
- f) Quite unintelligent use
- g) Extremely unintelligent use

The assumption is that the essential function of a competent leader is to distinguish between technically adequate and technically inadequate ideas. Finally, the conflicting suggestions by subjects were simply counted from a transcript of the conversation which was tape-recorded during the experimental sessions.

In addition to conflicting suggestions by the subjects, data from the transcript on the number of times the leader admitted failure and on the amount of opposition to the leader's suggestions and the number of negative remarks toward the leader—two possible manifestations of morale—are included in the report below.<sup>8</sup> Secondly, in the post-experimental questionnaire administered to each of the subjects, they were told that, if they considered dividing the cumulative earnings equally to be unfair, they could distribute their share among the other group members as they wished. The amount which they reassigned to the leader was taken as a measure of one aspect of morale. Finally, the following Coombs-type scales (4) were developed for the post-experimental questionnaire to measure other manifestations of morale:

Now that it's all over, how do you feel about participating in this group? Circle the alternative which best describes your feelings; "X" the alternative which is next best.

- a) I recall with enthusiasm the pleasure of participating with them
- b) I very much enjoyed the time spent participating with them
- c) I enjoyed the time I spent participating with them
- d) I enjoyed participating with them a little
- e) All in all, participating with them was slightly unattractive
- f) Participating with them was not attractive to me
- g) I found participating with them quite disagreeable

If you were to participate again in this group, how would you feel about having the same person serve as the group leader? Circle the best alternative; "X" the next best.

- a) Extremely favorable
- b) Quite favorable
- c) Somewhat favorable
- d) Not at all favorable
- e) Somewhat unfavorable
- f) Quite unfavorable
- g) Extremely unfavorable

Now taking into consideration who the leader would be, if you were to participate in this group again, how would you feel about having a leader make the decision again? Circle the best alternative; "X" the next best.

- a) Extremely favorable to a leader-type organization
- b) Quite favorable to a leader-type organization
- c) Somewhat favorable to a leader-type organization
- d) Not at all favorable to a leader-type organization
- e) Somewhat unfavorable to a leader-type organization
- f) Quite unfavorable to a leader-type organization
- g) Extremely unfavorable to a leader-type organization

<sup>8</sup> Only 4 of 20 interaction categories used in coding the transcriptions were included in this report, as the remainder were uncorrelated with the main variables of the study—perceived competence, reinforcement due to the leader, conflicting suggestions, and morale. As these variables involved indices of the leader's behavior, the lack of correlation indicates that the leader's behavior was consistent across the experimental variations. Inter-observer reliability, as indicated by agreement on items and not just totals for individuals or categories, exceeded .90.



How often did you feel yourself becoming angry with the leader? Circle the best alternative; "X" the next best.

- a) Continually
- b) Very frequently
- c) Quite frequently
- d) Once in a while
- e) Never

In other circumstances, how friendly would you be with the group leader? Circle the alternative which best describes your feelings; "X" the alternative which is next best.

- a) Extremely friendly
- b) Quite friendly
- c) Somewhat friendly
- d) Not friendly at all
- e) Somewhat unfriendly
- f) Quite unfriendly
- g) Extremely unfriendly

Thus, eight aspects of morale were measured in all. The scores for each measure were factor analyzed as part of the process of obtaining a measure of the underlying dimension. In addition to the underlying dimension of morale (Factor I), the factor analysis reported in Table 1 suggested the existence of two additional dimensions—expressive ambivalence (Factor II) and authoritarianism (Factor III). As all three appeared to be relevant to the study, indexes of each were calculated for each subject using standardized scores and the appropriate factor weights.

After the subjects completed the post-experimental questionnaire, they were informed about the details of the experiment. Interestingly, the deceptions worked in every case. However, cooling off the subjects usually

TABLE 1  
*Results of the Factor Analysis of the Morale Variables*

Variable	Factor Weights		
	I	II	III
1 Bonus to leader	.74	-.12	.03
2 Satisfaction with participation	.37	.32	.50
3 Satisfaction with leader	.81	.25	-.10
4 Satisfaction with leader-centered organization	.57	-.21	.46
5 Desire for friendship with leader	.42	.62	.18
6 Anger toward leader	-.61	-.22	.63
7 Opposition to leader's suggestions	-.34*	.67	-.28
8 Negative remarks	-.45	.63	.34

\* Only variables whose weights are equal to or greater than .35 are included in the calculation of a factor index score.

TABLE 2  
Matrix of Pearsonian Correlations

Variable	1	2	3	4	5	6	7	8
1 Relative reinforcement	—	.49	-.31	-.63	.57	-.36	.05	.98
2 Perceived competence		—	-.49	-.48	.61	-.05	.07	.47
3 Conflicting suggestions by subjects			—	.54	-.40	.38	.16	-.31
4 Admission of failure by leader				—	-.58	.22	-.01	-.59
5 Morale index					—	-.08	.07	.59
6 Expressive ambivalence index						—	-.04	-.38
7 Authoritarian index							—	.02
8 Success manipulation								—

necessitated rather lengthy discussion. But generally they responded well, showing considerable interest and quite voluntarily agreeing not to divulge any information about the experiment.

#### RESULTS

In Table 2 the intercorrelations between all of the variables measured in the experiment are reported. In passing it should be noted that the high, almost perfect correlation between success and relative reinforcement is an artifact of the experimental design. Rewards were assigned on the basis of the leader's success.

Indicating possible support of all three theories, the zero-order correlations between the three critical variables and morale are high and in the predicted direction. Relative reinforcement is correlated with morale .57, perceived competence with morale .61, and conflicting suggestions with morale  $-.40$ . All of these correlations are significant beyond the .05 level.<sup>9</sup>

<sup>9</sup> An explanation about the level of significance used in relation to the correlation coefficients is appropriate. It may be recalled that 21 groups and 42 subjects were used in this experiment. In order not to confound the effects of the variables such as authoritarianism, it was decided to calculate the coefficients using data for each of the 42 subjects, rather than the combined group data. This means that for some variables, such as production benefits from the leader and admission of failure by the leader, the subjects in the same groups have identical scores, having been exposed to identical stimuli. Although this may not be too critical since a case may be made for the statistical independence of the other variables, still, in statistics the conservative course is usually indicated. Consequently, in evaluating the significance of the correlation coefficients, an N of 21 was used rather than 42. Thus, to be significant at the .05 level, a zero-order correlation coefficient must be equal to or greater than .37, one-tailed test, and .43, two-tailed test.

However, reinforcement and perceived competence are significantly correlated, .49; and conflicting suggestions and perceived competence, -.49. Since these correlations are relatively large, it is possible that one or more of these three independent variables is spuriously correlated with morale. One of the variables—say, conflicting suggestions—may be correlated with morale simply because it is related to another variable that is more strongly correlated with morale—say, perceived competence. Consequently, an analysis restricted to the zero-order correlations is not sufficient.

To eliminate spurious relationships possibly contained in these intercorrelations, a multiple regression analysis can be used. First, the unique effect of each independent variable can be estimated by calculating partial regression coefficients, or beta weights. Secondly, the beta weights of each independent variable can be tested for significance. A significant beta weight is firm evidence that, with the effects of the other variables held constant, the relationship between the variable in question and morale is still significant. In addition, the multiple regression technique makes it possible to eliminate the effects of the tendency of the leader to admit failure—a tendency which, as we noted above, was not controlled experimentally. From Table 2 it is apparent that "admission of failure" is moderately associated with all of the critical variables in the study, and therefore should be controlled statistically.<sup>1</sup>

As may be noted from Table 3, when the multiple regression is calculated, it turns out that relative positive reinforcement independently controls 13 per cent of the variance of morale, perceived competence 23 per cent, conflicting suggestions 0 per cent, and admission of failure 15 per cent. Altogether, the four variables account for 51 per cent of the variance of morale; the multiple correlation of the four variables with morale is .72. *However, the tests of the beta weights indicate that only one of the four variables independently controls a significant proportion of the variance of morale.* The t-value of the beta weight for perceived competence is 1.73, which approximates the 1.75 required for significance at the 5 per cent level, using a conservative criterion. The other beta values do not reach significance using even a liberal criterion.<sup>10</sup> Thus, when the spurious effects are eliminated, the evidence supports but one of the three theories.

Two other relevant findings may be briefly summarized. First, the correlation of -.36 between relative reinforcement and expressive ambivalence—

<sup>10</sup> As discussed in footnote 9, such a criterion would use an N of 42 rather than 21, allowing 37 instead of 16 degrees of freedom. Using 37 degrees of freedom, the t-value for the beta value of reinforcement is 1.48; for perceived competence, 2.62; and for conflicting suggestions, 1.50. With 37 degrees of freedom, a t-value of 1.68 is required for significance at the .05 level.

TABLE 3

*Multiple Regression Correlation Analysis Between Four Variables and Morale*

Independent Variable	Beta Weight	t*	Explained Variance of Morale**
Reinforcement	-.23	1.00	.13
Perceived competence	.38	1.73	.23
Conflicting suggestions by subjects	-.01	—	.00
Admission of failure	-.25	1.00	.15

\* To be significant at the .05 level, the value of "t" must be equal to or greater than 1.75, using 16 degrees of freedom and a one-tailed test.

\*\* R is .72. The total explained variance of morale was partitioned among the four independent variables using the following formula:

$$R^2 = \beta_1 r_{01} + \beta_2 r_{02} + \beta_3 r_{03} + \beta_4 r_{04},$$

where the subscript "0" refers to morale. The four terms on the right-hand side of the equation represent the unique contribution of each independent variable. However, this is not the only way to partition explained variance. McNemar (14, p. 177) has suggested that the independent contribution of each variable is equal to the square of its beta weight. Calculated this way, the sum of the unique effects can exceed 100 per cent. Furthermore, this second approach necessitates additional terms, like  $\beta_1 \beta_2 r_{12}$ , which represent the common contribution of each pair of independent variables. These terms can be negative, indicating negative explained variance—a difficult concept to interpret. For both of these reasons, McNemar's approach was not used. A third approach is given by DuBois (5). Since the variance calculated by the DuBois method for each variable is not unique—it varies depending upon the order in which the variables are eliminated—this approach was not used either.

the tendency to show both positive and negative feelings—is of borderline significance. This indicates that relative reinforcement acts to suppress negative remarks and opposition to the leader, but, perhaps as a result, members like the leader less. While it would seem that expressive ambivalence should be associated with morale, the correlation is negligible. Secondly, an index of Factor III, which involves standard authoritarian reactions—consistent over- or under-reactions with respect to satisfaction with participation in a group where the decision was made by a leader, satisfaction with leader-centered organizations, and felt anger<sup>11</sup>—is not correlated significantly with any of the critical variables in this study. Thus, the relationship between morale and the perceived competence of the leader may occur quite independently of authoritarianism. However, in any future investigation of the morale-competence

<sup>11</sup> Note that this felt anger is toward the leader. Evidence from an experiment by Thibaut and Riecken (18) indicates that authoritarians do not have this overly-aggressive reaction toward leaders who have genuinely higher status. Presumably this over-reaction occurred because, in the present experiment, the leader was not of higher but of equal status—that of a student, the same as the participants.

hypothesis, it would be well to cross-check the validity of these last results with a standard measure.

#### DISCUSSION

In this investigation the results could have turned out in any one of several ways. The data could have supported none of the theories, any one of the theories, any two, or all three. However, as it is, the data support the perceived-competence theory, involving as it does the concept of reciprocity, but not the reinforcement nor the conflict theories.

In evaluating the results with respect to the three theories, it is critical to keep in mind that the measure of morale involves the reactions to many aspects of the organizational situation. In the reinforcement theory, an implicit assumption is that all aspects of an organization can be made acceptable simply by making the organization reinforcing or rewarding to the members. Certainly satisfaction or dissatisfaction with the rewards is important to most people in any organization, and satisfaction with rewards is expected to be related closely to relative rewards received. However, that this satisfaction with rewards should generalize to satisfaction with all aspects of the situation in the organization is a much more tenuous expectation. In an organization or in any other situation, is it ever possible, really, to buy friendship? Many, on the basis of their own personal observations, would say, "No! It is not possible!" The results with regard to expressive ambivalence in this experiment indicate that rewards did suppress overt disagreement and negative remarks. However, resentment also accrued which evidently emerged when the participants indicated they did not want to be friends with the leader. Such ambivalence may be a universal result of control by sheer reward, control by purchase.

Respect, acceptance, loyalty, and friendship, as reflected in morale, appear to be a function of organizational structure. Evidently organizational structure which is seen as violating the norm of reciprocity as it applies to assigned authority and perceived competence sets up a dissonance which destroys generalized satisfaction, including acceptance, respect, loyalty, and friendship.

Furthermore, it appears that the observed negative relationship between morale and conflicting suggestions is spurious. Evidently conflict itself did not reduce morale noticeably. The results of the multiple regression analysis indicate that it is not conflict per se that reduces morale, but perhaps conflict with the leader who is perceived as being incompetent. Thus, while a moderate negative correlation exists between perceived competence and conflicting suggestions, it is perceived competence that evidently influences morale and not the conflicting suggestions. These results suggest that conflict in and of itself is not necessarily harmful, but what is harmful is the way in which conflict is

resolved. Thus, it is not conflict that is to be avoided, but arbitrary resolutions of conflict which violate norms such as reciprocity. Since conflicting suggestions independently control none of the variance, we venture the opinion that the conflict theory is invalid.

One word of caution in interpreting the rest of the results: although this investigation does not support the reinforcement theory, the fact that reinforcement independently controlled 13 per cent of the variance of morale is important. Perhaps with a larger sample and with larger rewards over a longer period of time, a significant effect could be obtained. However, what is important here is the comparison in the reinforcement and perceived competence theories. Perceived competence not only controlled a significant amount of the variance, but controlled almost two times more of the variance of morale than did reinforcement. Thus, while this experiment may not be critical in the sense of eliminating the reinforcement theory, it may give a tentative comparison of the relative importance of the two theories.

That admission of failure by the leader should independently control 15 per cent of the variance of morale is, frankly, quite unexpected. Since the result could have occurred by chance, we shall not attempt an explanation. However, the morale-admission-of-failure hypothesis is of sufficient inherent interest to require further investigation.

Beyond this, the results of this investigation may have more general implications. Since, in most empirically-grounded theories of group behavior today, reinforcement is used exclusively in the derivation of predictions, such theories involve a psychologicistic approach to group behavior. The predictions based on these theories rely ultimately on the psychological properties of individuals, and this reliance may prevent a clear understanding of the functioning of organizations as social systems. By utilizing other bases of derivation involving structural variables such as the norm of reciprocity, we may uncover relationships that are not only significant but also important in terms of explained variance.<sup>12</sup>

<sup>12</sup> As this paper was going to press, it came to our attention that the phenomena described in this experiment might, for some purposes, be more adequately conceptualized in the framework suggested by Homans in his new book on social behavior (8). The reciprocity between morale and competence might be conceptualized in terms of distributive justice, and reinforcement in terms of reward. Homans suggests that sentiment (in our terms, morale) is a function of both optimizing reward and maintaining distributive justice. The results of this experiment indicate that distributive justice may be more important than reward. Homans also suggests the conditions under which the morale-competence hypothesis may not obtain, namely, under conditions where there is not a really obvious basis for establishing relative competence. Under such ambiguous conditions, length of service or tenure is often adopted as the main criterion for establishing rank. This usually does not solve the essential problem created by the ambiguities, however. Promotion by tenure probably does not erase conflicting evaluations of competence and,



## SUMMARY

The present investigation entailed an experiment designed to test and compare three competing theories involving independent derivations of the hypothesis that organizational morale varies directly with the relative competence of the leader. Reinforcement is the critical variable in the first theory, perceived competence the critical variable in the second, and conflicting suggestions by organizational members the critical variable in the third. While the observed correlations between the three critical variables and morale are significant and in the predicted direction, a multiple regression analysis supports but one of the three theories. Only perceived competence independently controlled a significant amount of the variance of morale.

*Manuscript received: July 14, 1960*

*Revised Manuscript received: May 8, 1961*

Robert L. Hamblin

The Social Science Institute

Washington University

St. Louis 30, Missouri

## REFERENCES

1. Baumgartel, H., "Leadership, Motivations, and Attitudes in Research Laboratories," *Journal of Social Issues*, 1956, 12, 24-31.
2. Coch, L., and J. R. P. French, Jr., "Overcoming Resistance to Change," *Human Relations*, 1948, 1, 512-532.
3. Cooley, C. H., *Social Organization*, Glencoe, Illinois: Free Press, 1956.
4. Coombs, C., "Theory and Methods of Social Measurement," in L. Festinger and D. Katz (eds.), *Research Methods in the Behavioral Sciences*, New York: Dryden, 1953.
5. DuBois, P. H., *Multivariate Correlation Analysis*, New York: Harper, 1957.
6. Durkheim, E., *The Division of Labor*, Glencoe, Illinois: Free Press, 1949.
7. Gouldner, A. W., "The Norm of Reciprocity: A Preliminary Statement," *American Sociological Review*, 1960, 25, 161-178.
8. Homans, G. K., *Social Behavior: Its Elementary Forms*, New York: Harcourt, Brace, 1961.
9. Katz, D., and R. Kahn, "Some Recent Findings in Human Relations Research in Industry," in G. Swanson, T. Newcomb, and E. Hartley (eds.), *Readings in Social Psychology*, New York: Holt, 1952.
10. Leavitt, H. J., "Some Effects of Certain Communication Patterns on Group Performance," *Journal of Abnormal and Social Psychology*, 1951, 46, 38-50.

---

consequently, does not repair the harm to morale which evidently results when the perceived competence of those of higher rank is relatively low. Under ambiguous conditions, a more effective solution would involve the introduction and use of objective measures of competence such as the fleet-wide tests introduced and used by the U. S. Navy.

11. Lewin, K., and R. Lippitt, "An Experimental Approach to the Study of Autocracy and Democracy: A Note," *Sociometry*, 1938, 1, 292-300.
12. Lewin, K., R. Lippitt, and R. K. White, "Patterns of Aggressive Behavior in Experimentally Created 'Social Climates,'" *Journal of Social Psychology*, 1939, 10, 271-299.
13. Maier, N. R. F., "The Quality of Group Decisions as Influenced by the Discussion Leader," *Human Relations*, 1950, 3, 155-174.
14. McNemar, Q., *Psychological Statistics*, New York: Wiley, 1955.
15. Merton, R. K., and A. S. Kitt, "Contributions to the Theory of Reference-Group Behavior," in G. Swanson, T. Newcomb, and E. Hartley (eds.), *Readings in Social Psychology*, New York: Holt, 1952.
16. Sanford, F. H., *Authoritarianism and Leadership*, Philadelphia: Institute for Research in Human Relations, 1950.
17. Simmel, G., *The Sociology of Georg Simmel*, edited by K. Wolff, Glencoe, Illinois: Free Press, 1950.
18. Thibaut, J., and H. Riecken, "Authoritarianism, Status, and the Communication of Aggression," *Human Relations*, 1955, 8, 95-120.
19. Tinklepaugh, O. L., "An Experimental Study of Representative Factors in Monkeys," *Journal of Comparative Psychology*, 1928, 8, 187-236.

## A Note on Statistical Significance of Scalogram<sup>1</sup>

KARL F. SCHUESSLER, *Indiana University*

In his recent book on measurement, W. S. Torgerson states that one limitation of Guttman's method of scaling resides in its incapacity to test statistically whether the cumulative scale model is a good fit (2). The Guttman scale model (1), it will be recalled, consists of an ordered set of  $\left[ \left( \sum_1^m c_i - m \right) + 1 \right]$  combinations from a total of  $(c_1 c_2 \dots c_i \dots c_m)$  possible combinations, where  $m$  is the number of items, and  $c_i$  is the number of categories of the  $i$ th item. It is by definition a single-valued function in which each scale score,  $x$ , corresponds to only one combination,  $y$ .

While there is truth to Torgerson's statement, it is perhaps unnecessarily categorical and may act to divert attention from scale analysis before its potentialities as an ordering device have been properly considered by the investigator. For it is possible to construe the matrix of observations (scalogram) in such a way that statistical hypotheses may be tested, and inferences drawn in respect to the presence of basic dimensions which could account for the configuration of observed responses. This paper sets forth three such possibilities, primarily as suggestions to the research worker who is uncertain about the statistical significance of his scale findings and who wishes as a minimum to rule out chance as an explanation.

Although the methods to be unfolded are general, as a convenience we shall have recourse throughout to the hypothetical responses of 100 individuals to three dichotomized attitude items, with the following marginals:

Item	Response *	
	+	-
I	.80	.20
II	.60	.40
III	.35	.65

\* Plus sign represents positive response; minus sign, a negative response.

<sup>1</sup> The author is indebted to Professor Louis Guttman of Hebrew University, Jerusalem, Israel, who kindly read the original manuscript of this note and made suggestions for its refinement.

and the following distribution of persons by response patterns:

Number of Positive Responses	Response Pattern			Frequency
( $x$ )	I	II	III	
3	+	+	+	30
2	+	+	—	25
2	+	—	+	3
2	—	+	+	1
1	+	—	—	22
1	—	+	—	4
1	—	—	+	1
0	—	—	—	14
				$n = 100$

Finally, for ease of exposition, we shall employ, along with  $x$ ,  $m$ , and  $c_i$ , the following notation:

$n$  = number of persons in sample

$n_{ij}$  = number of persons giving  $j$ th response to  $i$ th item

$p_{ij}$  = proportion of persons giving  $j$ th response to  $i$ th item

#### TEST I

Essentially, this procedure is a comparison between the number of scale types expected on the hypothesis that responses to successive items are statistically independent and the number of scale types observed in a sample of  $n$  persons. If this difference is sufficiently improbable by chi-square, then the hypothesis of independence may be waived and an alternative hypothesis accepted, whether or not explicitly formulated. Test I, then, takes as its point of departure the proportion of scale types expected by chance, against which the observed proportion of scale types is evaluated for possible significance.

As a preliminary step, we must designate scale types, since it is their combined frequency that is to be checked for statistical significance. In general, we designate as scale types those arrangements which have the largest chance expectation within their respective score intervals, and which together permit a perfect fit. In the event of ties within a given score interval, we designate whichever arrangement has the largest observed frequency, since that arrangement will yield a better fit than any other combination in the same score class.

To illustrate the identification of scale types by this method, we calculate on page 314 the chance frequencies corresponding to the response patterns given above:

$s$	Combination	$p_{11}$	$p_{21}$	$p_{31}$	$(p_{11}p_{21}p_{31})$	Expected Frequency
*3	+ + +	.80	.60	.35	.168	16.8
*2	+ + -	.80	.60	.65	.312	31.2
2	+ - +	.80	.40	.35	.112	11.2
2	- + +	.20	.60	.35	.042	4.2
*1	+ - -	.80	.40	.65	.208	20.8
1	- + -	.20	.60	.65	.078	7.8
1	- - +	.20	.40	.35	.028	2.8
*0	- - -	.20	.40	.65	.052	5.2
					1.000	100.0

\* Denotes scale type.

Of those arrangements yielding a scale score of two, we treat  $++-$  as the scale type, since it has the largest chance frequency (31.2); similarly, of the arrangements producing a score of one, we designate  $+--$  as the scale type, since its chance expectation of 20.8 is largest. Where a given score can result from only one arrangement, as in the case of extreme scores, no choice is possible, so that extreme scores, when they occur, must always be designated as scale types.

Once scale types have been identified, we merely add their individual probabilities,\* and treat the resulting sum as the probability of a scale type on a single trial. We next apply this probability to  $n$  to obtain the frequency of scale types expected on the hypothesis that the  $m$  items are statistically independent. Lastly, against this expected frequency we check the observed frequency for possible statistical significance. Combining scale types and their corresponding probabilities, we obtain:

$$\begin{aligned} Pr(\text{Scale Type}) &= .168 + .312 + .208 + .052 \\ &= .74 \end{aligned}$$

Accordingly, the expected frequency of scale types is  $(.74)(100) = 74$ , which leads to  $\chi^2 = 15.02$ , 1 *df*. Now, since we are interested only in the possibility that the observed frequency of scale types exceeds the frequency expected by chance, and not in the possibility of its being deficient, we take the positive square root of 15.02 and calculate the probability of  $z$  (normal deviate) larger than  $15.02 = 3.87$ . Since the probability value in this case is very small, we abandon the initial hypothesis that the  $m$  items are uncorrelated and accept the possibility that determining factors are present.

#### ALTERNATIVE HYPOTHESES

At this juncture, it would be possible to give consideration to alternative

hypotheses of varying degrees of stringency. For instance, we might hypothesize that the probability of a scale type is equal to, or less than, .8 ( $H: P \leq .8$ ), and proceed to determine the probability of the observed frequency of scale types against that hypothesis. Since any frequency smaller than  $.8n$  is consistent with this one-tailed hypothesis, we need consider only those observed frequencies which are larger than  $.8n$ , as, for example,  $.91n = 91$ , in our illustration. Upon testing 91 against the expected frequency of 80, we obtain  $\chi^2 = 7.44$ ,  $z = 2.72$ , and  $Pr < .01$ , a significance value which would prompt rejection of  $H: P \leq .8$ , and acceptance of the alternative that the probability of a scale type is greater than .8.

By giving due consideration to such alternatives to the chance hypothesis, we meet in part Torgerson's comment that scale analysis cannot be checked for goodness of fit. While it is not possible to test the hypothesis of perfect scalability by probabilistic methods, still we may test weaker alternatives, including the chance hypothesis, and by that technique assure ourselves, as a minimum, that the chance hypothesis is probably false. More positively, we may assess the tendency in the sampled population to adhere to the scale model, although we can never prove that the population is perfectly scalable.

#### TEST II

Like the first, this procedure requires that we initially calculate the chance frequencies of the respective response patterns, and designate scale types accordingly. But, instead of combining scale types into a single class, we now compare the observed and expected frequency of scale types within specific score intervals, along with the observed and expected frequency of non-scale types.<sup>2</sup> Obviously, this procedure is more sensitive than the first, since the comparisons within specific score intervals will indicate at what points the divergence from chance is negligible, and thereby lead to a more refined statistical interpretation of the matrix of responses. However, since  $\chi^2$  will necessarily be based on more than 1 *df*, it will be impossible to test by this statistic the one-tailed directional hypothesis that the observed scale frequencies are excessive. It will be possible to test only the hypothesis that scale frequencies are a chance phenomenon. Consequently, it will be necessary to determine visually whether the observed scale frequencies exceed, or fall short of, those expected by chance. This procedure is illustrated in the tabulation on page 316:

<sup>2</sup> Since each score yields two comparisons, excepting the extremes which yield one, the total number of comparisons will be  $2(\Sigma c_i - m)$ .



$x$	$O$	$E$	$(O-E)$	$\frac{(O-E)^2}{E}$
3	30	16.8	13.2	10.37
2	25	31.2	-6.2	1.24
2	4	15.4	-11.4	8.44
1	22	20.8	1.2	.07
1	5	10.6	-5.6	2.96
0	14	5.2	8.8	14.90
	<hr/> 100	<hr/> 100.0	<hr/> 0.0	<hr/> 37.98

Inasmuch as  $Pr(\chi^2 > 37.98/5 \text{ df}) < .01$ , we reject the hypothesis that items are independent and tacitly accept, as in Test I, the hypothesis of response dependency.

### TEST III

As a third possibility, we may analyze the variation between and within items for each set of persons having the same score, a procedure which will usually result in  $[(\sum c_i - m) - 1]$  separate analyses, or two less than the number of scale types. If persons with identical scores are consistently alike in their response to each item, then the total response variation will be completely accounted for by whatever variation exists between items and the intraclass correlation coefficient,  $\delta$ , will be 1. At the other extreme, where total response variation is almost wholly accounted for by the within-item variation, the intraclass correlation coefficient will generally lie close to zero, thereby demonstrating a random distribution of values within classes. Evidently, we may employ the intraclass correlation coefficient to test the hypothesis that items are uncorrelated within each subset of identical scores, and to gauge the degree of score homogeneity within classes.

To exemplify this type of analysis, let us return once again to the table above, in which 29 persons show scores of two, and 27 persons show scores of one. (Extreme scores are not subject to this analysis since there is no variation to be partitioned.) For each group, we set up a person-by-item matrix, the elements of which are arbitrary values—0 or 1 in our example—and proceed to the calculation of the intraclass coefficient (computation details available on request to the author). The  $29 \times 3$  score matrix yields  $\delta = .71$ , and the  $27 \times 3$  matrix yields  $\delta = .61$ . Both results are sufficiently improbable by the F-statistic to warrant rejection of the null hypothesis under test.

### COMPARISON OF PROCEDURES

While the techniques presented above will normally yield identical decisions, as in our running illustration, they may at times lead to conflicting inferences

(rejection *and* acceptance), owing to the fact that they answer to different but related questions and/or rest on different assumptions concerning the observed data.

Procedures I and II are essentially alike in testing an observed frequency against an expected frequency, but they differ in that I merges all scale types together before testing for what may be called gross significance, while II tests the observed frequency of scale types against the expected frequency within specific score intervals.

Both I and II assume that responses are purely qualitative, whereas III treats them as numerical values subject to arithmetic averaging. It follows that conflicting decisions between II and III, both of which examine the pattern of responses within score intervals, could result from the fact that in one instance observations are arbitrarily quantified, whereas in the other case they are merely enumerated. This difference points up a possible limitation of III, and a corresponding advantage of II: the former has recourse to artificial measures, while the latter only counts events.

Further, the intraclass correlation coefficient is a function of the variance ratio, and, consequently, to test its significance it is necessary to satisfy the conditions of the F-distribution. On the other hand, it is a conventional measure of correlation, and if such a measure be required, then its use would be dictated by that need, even though its statistical significance could not be exactly determined. But, all in all, the more conservative distribution-free method is less open to question.

#### REPRODUCIBILITY AND CHANCE FREQUENCY OF SCALE TYPES

Since the procedures suggested in this paper intend primarily to rule out chance as an explanation of the scalogram, they in no way supplant Guttman's well-known criteria of a perfect scale. They would normally be applied in advance of Guttman's four principal criteria, namely: (a) 90 per cent reproducibility, (b) balanced marginals, (c) less error than non-error within items, and (d) random scatter of error within items. At times, to be sure, the significance of the scalogram may appear obvious and a null test of chance unnecessary. But, whenever doubt exists within the mind of the worker, then null tests should be run as a matter of course. In particular, such checks are essential when the number of scale types is small [ $(\sum c_i - m + 1) < 10$ ] and/or the marginals are uniformly unbalanced (e.g.,  $p_{ij} > .80$ ). For in those situations, the frequency of scale types will necessarily be high as a matter of chance, and is likely therefore to be spuriously credited with significance when none actually exists.

It is evident from inspection of the foregoing examples that reproducibility need not be calculated in order to apply any of the foregoing tests. However,

each procedure automatically provides a test of the hypothesis that reproducibility is a chance outcome, since reproducibility is a function of the ratio of scale types to non-scale types, increasing as that ratio increases. Hence, a demonstration of the significance of this ratio is also a demonstration of the significance of observed reproducibility. But since the number of scale types expected by chance is much simpler to calculate than chance reproducibility, we naturally have recourse to that simpler quantity.

#### SUMMARY

It is the thesis of this paper that Guttman scalograms may be statistically interpreted and that such interpretations have a bearing on the decision to accept the hypothesis that the items constitute a scale or quasi-scale. While the Guttman hypothesis of perfect scalability is statistically untestable, it is possible to determine whether the observed configuration of responses might have arisen by chance. Additionally, alternative hypotheses may be tested in order to set a probable lower boundary to the proportion of scale types in the sampled population and thereby to gauge the tendency in the population to conform to the scale model. Altogether, the methods set forth in this paper enable the research worker to approach his scale findings within the framework of statistical inference, and correspondingly to make his decisions with greater or lesser confidence.

*Manuscript received: August 19, 1960*

*Revised Manuscript received: December 20, 1960*

*Karl F. Schuessler*

*Department of Sociology*

*Indiana University*

*Bloomington, Indiana*

#### REFERENCES

1. Guttman, L., "The Quantification of a Class of Attributes," in Paul Horst (ed.), *The Prediction of Personal Adjustment*, Social Science Research Council, Bulletin 48, 1941, 319-348.
2. Torgerson, W. S., *Theory and Methods of Scaling*, New York: John Wiley & Sons, 1958, 324.

